77501

STIC-Biotech/ChemLib

From: Helmer, Georgia
Sent: Tuesday, October 08, 2002 6:17 PM
To: STIC-Biotech/ChemLib
Subject: Seq Search - 09/289,346

a 42-55

In case - 09/289,346:

could you please do a sequence search and an oligo search on SEQ ID Nos 1 thru 10.

Both inhouse and commercial databases. And also an Interference search.

Thank you in advance for your help.

Georgia L. Helmer Ph.D.
Patent Examiner
Crystal Mall 1, 9D14
AU 1638
703-308-7023
Georgia.Helmer@USPTO.gov
mailbox 9e12

seq ID#1 = wt

RECEIVED
OCT - 9 2002
(STIC)

Point of Contact
P. Sheppard
Telephone number: (703) 308-4499

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 10/11/02
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

Thu Oct 10 16:05:31 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:43:47 ; Search time 118.38 Seconds
(without alignments)
102.295 Million cell updates/sec

Title: US-09-289-346A-1
Perfect score: 364
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	315	86.5	351	12	Q91R10	Q91r10 tomato seve
2	313	86.0	352	12	Q9E000	Q9e000 tomato rugo
3	296	81.3	226	12	O09727	O09727 leonurus mo
4	296	81.3	226	12	Q9WHF6	Q9whf6 tomato mild
5	294	80.8	361	12	Q67574	Q67574 bean golden
6	292	80.2	225	12	Q9QDB1	Q9qdb1 cowpea gold
7	292	80.2	314	12	Q9ELT8	Q9elt8 sweet potat
8	292	80.2	364	12	Q9QS55	Q9qs55 sweet potat
9	289	79.4	185	12	Q98693	Q98693 sida golden
10	281	77.2	149	12	P88975	P88975 macroptiliu
11	281	77.2	233	12	Q9YLA4	Q9yla4 macroptiliu
12	275	75.5	234	12	O39180	O39180 geminivirid
13	272	74.7	190	12	Q9Z089	Q9z089 tobacco lea
14	272	74.7	190	12	Q9Z084	Q9z084 tobacco lea
15	271	74.5	190	12	Q9W827	Q9w827 tobacco lea
16	271	74.5	208	12	Q9Z0C4	Q9z0c4 tobacco lea
17	271	74.5	359	12	Q91M88	Q91m88 tobacco lea
18	269	73.9	208	12	Q9Z0C0	Q9z0c0 tobacco lea
19	269	73.9	208	12	Q9Z0B8	Q9z0b8 tobacco lea
20	267	73.4	208	12	Q9Z0B6	Q9z0b6 tobacco lea
21	266	73.1	203	12	Q9Z083	Q9z083 tobacco lea
22	266	73.1	363	12	O73577	O73577 cotton leaf
23	261	71.7	359	12	Q91B86	Q91b86 ageratum ye
24	261	71.7	360	12	Q9DXI0	Q9dxi0 ageratum ye
25	259	71.2	208	12	Q9Z0A0	Q9z0a0 tobacco lea
26	259	71.2	363	12	O72705	O72705 cotton leaf
27	259	71.2	363	12	O72719	O72719 cotton leaf
28	258	70.9	190	12	Q9Z086	Q9z086 tobacco lea
29	258	70.9	208	12	Q9Z0C6	Q9z0c6 tobacco lea
30	257	70.6	349	12	Q88888	Q88888 tomato pseu
31	257	70.6	362	12	O56816	O56816 chayote mos
32	256	70.3	359	12	Q9YZV4	Q9yzv4 tomato yell
33	256	70.3	359	12	Q9YZV2	Q9yzv2 tomato yell
34	256	70.3	359	12	Q9YUX7	Q9yux7 tomato yell
35	256	70.3	359	12	Q9YL27	Q9yl27 tomato yell
36	255	70.1	348	12	Q911W5	Q911w5 macroptiliu
37	255	70.1	353	12	O72692	O72692 beet curly
38	255	70.1	354	12	Q65438	Q65438 beet curly
39	255	70.1	359	12	Q88942	Q88942 tomato yell
40	254	69.8	190	12	Q9Z0A7	Q9z0a7 tobacco lea
41	254	69.8	362	12	Q9IN48	Q9in48 okra enatio
42	253	69.5	307	12	Q9IET7	Q9iet7 cotton leaf
43	253	69.5	361	12	O72723	O72723 cotton leaf
44	252	69.2	231	12	Q96620	Q96620 african tom
45	252	69.2	354	12	Q9IN42	Q9in42 south afric

ALIGNMENTS

RESULT 1
Q91R10 PRELIMINARY; PRT; 351 AA.
ID Q91R10
AC Q91R10;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE REP PROTEIN.
GN AC1.
OS Tomato severe rugose virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=158463;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MINAS GERAIS;
RA Rezende W.L., Goulart L.R., Parreira K.S., Figueiredo J.E.F.;
RT "The full-length DNA-A nucleotide sequence of a novel tomato-infecting
RT begomovirus, Tomato severe rugose virus, in Brazil.";
RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY029750; AAK50357.1; -.
SQ SEQUENCE 351 AA; 40122 MW; 87F937A4F873B6CF CRC64;

Query Match 86.5%; Score 315; DB 12; Length 351;
Best Local Similarity 82.9%; Pred. No. 2.9e-27;
Matches 58; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||||:|||||||||||:|||||||||| ||: |||||||:|||:|||||||||||
Db     111 TIEWGEFQIDGRSARGGCQTANDAAAEALNAPSKDVALQIIREKLPEKFLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| : |||
Db     171 DRIFARAPEP 180
```

RESULT 2
Q9E000 PRELIMINARY; PRT; 352 AA.
ID Q9E000

```
AC   Q9E000;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN.
GN   AC1.
OS   Tomato rugose mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=134599;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Fernandes J.J., Fontes E.P.B., Brommonschenkel S.H., Carvalho M.G.,
RA   Zambolim E.M., Zerbini F.M.;
RT   "Molecular Cloning and Characterization of Tomato rugose mosaic virus
RT   (TRMV), a Begomovirus Isolated from Tomatoes at Triangulo Mineiro,
RT   Minas Gerais, Brazil.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF291705; AAG15546.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   352 AA;  40012 MW;  47CD55838E24D613 CRC64;

  Query Match                    86.0%;  Score 313;  DB 12;  Length 352;
  Best Local Similarity   82.9%;  Pred. No. 4.9e-27;
  Matches   58;  Conservative    7;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||||||:|||||||||| ||: |||||||:|||:||||||||||
Db     111 TIEWGEFQIDGRSARGGCQTANDAAAEALNAPSKDIALQIIREKMPEKFLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| : |||
Db     171 DRIFARAPEP 180


RESULT   3
O09727
ID   O09727
AC   O09727;      PRELIMINARY;      PRT;   226 AA.
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   REP.
OS   Leonurus mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=58177;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LEMV- BRAZIL 1;
RA   Faria J.C., Maxwell D.P.;
RT   "Variability in geminivirus associated with Phaseolus vulgaris in
RT   Brazil.";
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U92532; AAB51157.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     226     226
SQ   SEQUENCE   226 AA;  25617 MW;  73CDB6E766083FC5 CRC64;

  Query Match                    81.3%;  Score 296;  DB 12;  Length 226;
  Best Local Similarity   81.4%;  Pred. No. 2.4e-25;
  Matches   57;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||||||||||:||| || |||||||||| |  ||||||:|||:||||||||||||||||
Db     111 TVEWGEFQVDGRSSRGGQQTVNDAAAEALNAPDKRTALQIIKEKLPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180


RESULT   4
Q9WHF6
ID   Q9WHF6
AC   Q9WHF6;      PRELIMINARY;      PRT;   226 AA.
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   tomato mild mottle geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX   NCBI_TaxID=92943;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HN96-H5KW;
RA   Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA   Maxwell D.P.;
RT   "Molecular characterization and DNA-based detection methods for
RT   vegetable-infecting geminiviruses in Central America.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF131071; AAD33471.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     226     226
SQ   SEQUENCE   226 AA;  25941 MW;  2EA4116712871A23 CRC64;

  Query Match                    81.3%;  Score 296;  DB 12;  Length 226;
  Best Local Similarity   77.1%;  Pred. No. 2.4e-25;
  Matches   54;  Conservative   12;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| ||:|||||||||||||||||:||:||:|:::|||:|||:|||
Db     111 TIEWGEFQIDGRSARGGQQTANDAAAEALNASSKEEAMRIIKEKLPKEFLFQYHNLSSNL 170

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180


RESULT   5
Q67574
ID   Q67574      PRELIMINARY;      PRT;   361 AA.
AC   Q67574;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE  REPLICATIVE PROTEIN.
GN   AL1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,
RA   Maxwell D.P., Russell D.R.;
RT   "Cloning of the complete DNA genomes of four bean-infecting
RT   geminiviruses and determining their infectivity by electric discharge
RT   particle acceleration.";
RL   Phytopathology 81:980-985(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,
RA   Morales F.J., Maxwell D.P.;
RT   "Differentiation of bean-infecting geminiviruses by nucleic acid
```

```
RT   hybridization probes and aspects of bean golden mosaic in Brazil.";
RL   Plant Dis. 75:336-342(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RT   "Genetic diversity in geminiviruses causing bean golden mosaic
RT   disease: The nucleotide sequence of the infectious cloned DNA
RT   components of a Brazilian isolate of bean golden mosaic geminivirus.";
RL   Submitted (MAY-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; M88686; AAA46312.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   361 AA;  41041 MW;  0094C7ACAF06B788 CRC64;

  Query Match             80.8%;  Score 294;  DB 12;  Length 361;
  Best Local Similarity   80.6%;  Pred. No. 6.8e-25;
  Matches   54;  Conservative    9;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           || |||||||||||| ||:||||:||||||||||:|||:|||:|||:|||:||||||
Db     113 WGHFQVDGRSARGGQQTANDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
             | | |:|
Db     173 FTKAPDP 179


RESULT   6
Q9QDB1
ID   Q9QDB1      PRELIMINARY;      PRT;   225 AA.
AC   Q9QDB1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   cowpea golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=69263;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CGMV-BR;
RA   Faria J.C.;
RT   "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT   Brazil.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF188708; AAF06318.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     225    225
SQ   SEQUENCE   225 AA;  25766 MW;  1089CB6BD8D15B5D CRC64;

  Query Match             80.2%;  Score 292;  DB 12;  Length 225;
  Best Local Similarity   80.6%;  Pred. No. 6.6e-25;
  Matches   54;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           || ||:||||||| || ||||:|||||||||||:|||:||:|||:|||:|||:||||||
Db     113 WGHFQIDGRSARGGQQTINDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
             | | |||
Db     173 FKKPPEP 179


RESULT   7
Q9ELT8
ID   Q9ELT8      PRELIMINARY;      PRT;   314 AA.
AC   Q9ELT8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION ASSCCIATION PROTEIN.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF288227; AAG01006.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   314 AA;  35153 MW;  6B6220613046943F CRC64;

  Query Match             80.2%;  Score 292;  DB 12;  Length 314;
  Best Local Similarity   83.8%;  Pred. No. 9.7e-25;
  Matches   57;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||||||||||| ||:|||||||||| ||| ||||||||:|||||||||||| |||
Db     110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy      61 DRIFDKTP 68
           ||||   |
Db     170 DRIFSPPP 177


RESULT   8
Q9QS55
ID   Q9QS55      PRELIMINARY;      PRT;   364 AA.
AC   Q9QS55;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION INITIATION PROTEIN AC1.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A.;
RT   "Cloning of a DNA-A-like genomic component of sweet potato leaf curl
RT   virus : nucleotide sequence and phylogenetic relationships.";
RL   Online Publication.
DR   EMBL; AF104036; AAD47173.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   364 AA;  40680 MW;  5F79752431A09D6E CRC64;
```

```
  Query Match             80.2%;  Score 292;  DB 12;  Length 364;
  Best Local Similarity   83.8%;  Pred. No. 1.1e-24;
  Matches   57;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||||||||||||||| ||:|||||||||| ||| ||||||||:|||||||||| |||
Db     110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy      61 DRIFDKTP 68
           ||||   |
Db     170 DRIFSPPP 177


RESULT   9
Q98693
ID   Q98693       PRELIMINARY;      PRT;   185 AA.
AC   Q98693;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   AC1.
OS   sida golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51034;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA;
RA   Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RT   "Genetic Diversity among geminiviruses associated with the weed
RT   species Sida spp, Macroptilium lathyroides, and Wissadula amplissima
RT   from Jamaica.";
RL   Plant Dis. 81:1251-1258(1997).
DR   EMBL; U67926; AAB97865.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     185    185
SQ   SEQUENCE   185 AA;  20975 MW;  3913850A025A5EE1 CRC64;


  Query Match             79.4%;  Score 289;  DB 12;  Length 185;
  Best Local Similarity   74.3%;  Pred. No. 1.1e-24;
  Matches   52;  Conservative   12;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:|||||||| ||:|||||||||: :||:||:|||||:||:||||:|||:|||:
Db      89 TIEWGVFQIDGRSARGGQQTANDAAAEALNSGTKEDALKIIREKLPERYLFQYHNLSSNI 148

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     149 DRIFSKPPEP 158


RESULT  10
P88975
ID   P88975       PRELIMINARY;      PRT;   149 AA.
AC   P88975;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   AC1.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICAN;
RA   Roye M.E.;
RL   Thesis (1996), Biochemistry, University of the West Indies, Jamaica.
DR   EMBL; U75278; AAB36919.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     149    149
SQ   SEQUENCE   149 AA;  16785 MW;  E4CF5EED4C9CD508 CRC64;


  Query Match             77.2%;  Score 281;  DB 12;  Length 149;
  Best Local Similarity   72.9%;  Pred. No. 7.1e-24;
  Matches   51;  Conservative   12;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:|||||||| ||||||||||||: :|| |::|::||:|||:|||:|||:|||
Db      52 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 111

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     112 DRIFMKDPEP 121


RESULT  11
Q9YLA4
ID   Q9YLA4       PRELIMINARY;      PRT;   233 AA.
AC   Q9YLA4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E.;
RT   "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
RT   from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E., McLaughlin W.A., Maxwell D.P.;
RT   "Molecular characterization of two distinct geminiviruses infecting M.
RT   lathyroides from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF098940; AAD17850.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     233    233
SQ   SEQUENCE   233 AA;  26356 MW;  AA490AF4D2166A02 CRC64;


  Query Match             77.2%;  Score 281;  DB 12;  Length 233;
  Best Local Similarity   72.9%;  Pred. No. 1.2e-23;
  Matches   51;  Conservative   12;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:|||||||| ||||||||||||: :|| |::|::||:|||:|||:|||:|||
Db     110 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 169

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     170 DRIFMKDPEP 179


RESULT  12
O39180
```

```
ID   O39180      PRELIMINARY;      PRT;   234 AA.
AC   O39180;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   AC1 (FRAGMENT).
OS   Geminiviridae.
OC   Viruses; ssDNA viruses.
OX   NCBI_TaxID=10811;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOMATO STRAIN;
RA   Guzman P., Arredondo C.R., Emmatty D., Portillo R.J., Gilbertson R.L.;
RT   "Partial Characterization of Two Whitefly-Transmitted Geminiviruses
RT   Infecting Tomatoes in Venezuela.";
RL   Plant Dis. 81:312-312(1997).
DR   EMBL; AF026553; AAB82605.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     234    234
SQ   SEQUENCE   234 AA;  26486 MW;  9ED8F0697105CD19 CRC64;

  Query Match             75.5%;  Score 275;  DB 12;  Length 234;
  Best Local Similarity   71.4%;  Pred. No. 5.6e-23;
  Matches   50;  Conservative   13;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:|||||||| ||:|||||||||: :||||::||:|:|:|||:|||:|||:|:|:|
Db     110 TIEWGLFQIDGRSARGGQQTANDAAAEALNSGTKEEAMKIIKEKLPEKFLFQYHNLSSDL 169

Qy      61 DRIFDKTPEP 70
           |||| | | |
Db     170 DRIFMKAPNP 179


RESULT  13
Q9Z089
ID   Q9Z089      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z089;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA3-1, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001315; BAA34033.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21432 MW;  AAC093D1D1610FAD CRC64;

  Query Match             74.7%;  Score 272;  DB 12;  Length 190;
  Best Local Similarity   62.4%;  Pred. No. 9.5e-23;
  Matches   53;  Conservative    8;  Mismatches    8;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || || ||:|||||||||| :||| ||||||||| ||| |||||:|: ::||:|||||||
Db      85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  14
Q9Z084
ID   Q9Z084      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z084;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA5-2, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA5;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001318; BAA34039.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  AAC1C2943E3F01AD CRC64;

  Query Match             74.7%;  Score 272;  DB 12;  Length 190;
  Best Local Similarity   62.4%;  Pred. No. 9.5e-23;
  Matches   53;  Conservative    8;  Mismatches    8;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || || ||:|||||||||| :||| ||||||||| ||| |||||:|: ::||:|||||||
Db      85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  15
Q9W827
ID   Q9W827      PRELIMINARY;      PRT;   190 AA.
AC   Q9W827;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE GORA1-1, PARTIAL AND
DE   COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GORA1;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001303; BAA34010.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
```

```
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  93C3742A8EBDB7EB CRC64;


  Query Match             74.5%;  Score 271;  DB 12;  Length 190;
  Best Local Similarity   62.4%;  Pred. No. 1.2e-22;
  Matches   53;  Conservative    8;  Mismatches    8;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || || |||||||||||||| :||| |||||||:|| ||| |||||:|: ::||:|||||||
Db      85 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFSNSSFDQVPE 169


Search completed: October 10, 2002, 02:01:59
Job time: 1092 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:47:22 ; Search time 34.62 Seconds
(without alignments)
78.289 Million cell updates/sec

Title: US-09-289-346A-1
Perfect score: 364
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	364	100.0	352	1	VAL1_TGMV	P03567 tomato gold
2	264	72.5	361	1	VAL1_PYMVV	P27258 potato yell
3	251	69.0	358	1	VAL1_CLVK	P14982 cassava lat
4	251	69.0	358	1	VAL1_CLVN	P14972 cassava lat
5	244	67.0	362	1	VAL1_TYLCA	P36279 tomato yell
6	237	65.1	359	1	VAL1_TYLCU	P38609 tomato yell
7	233	64.0	349	1	VAL1_PHUV	Q06923 pepper huas
8	230	63.2	359	1	VAL1_TYLCM	P27260 tomato yell
9	228	62.6	353	1	VAL1_BGMV	P05175 bean golden
10	227	62.4	355	1	VAL1_ABMVW	P21947 abutilon mo
11	224	61.5	358	1	VAL1_BCTV	P14991 beet curly
12	224	61.5	361	1	VAL1_TMOV	Q06657 tomato mott
13	215	59.1	357	1	VAL1_TYLCV	P27259 tomato yell
14	124	34.1	347	1	VAL1_SLCV	P29048 squash leaf
15	64.5	17.7	299	1	Y175_HELPJ	Q9zmq7 helicobacte
16	64.5	17.7	1713	1	LMA3_HUMAN	Q16787 homo sapien
17	63.5	17.4	1610	1	CCAD_MESAU	Q99244 mesocricetu
18	63.5	17.4	2161	1	CCAD_HUMAN	Q01668 homo sapien
19	63.5	17.4	2203	1	CCAD_RAT	P27732 rattus norv
20	62.5	17.2	447	1	TBL2_HUMAN	Q9y4p3 homo sapien
21	62	17.0	295	1	VAL1_TYDVA	P31617 tobacco yel
22	61	16.8	1483	1	CYP1_YEAST	P12351 saccharomyc
23	60	16.5	447	1	CDSA_DROME	P56079 d phosphati
24	59	16.2	387	1	Y4PF_RHISN	P55615 rhizobium s
25	58.5	16.1	136	1	Y452_CAEEL	O62250 caenorhabdi
26	58.5	16.1	297	1	RRPP_RABVP	P06747 rabies viru
27	58.5	16.1	299	1	Y175_HELPY	P56112 helicobacte
28	58.5	16.1	428	1	GBA1_CANAL	P28868 candida alb
29	58.5	16.1	589	1	YE67_HAEIN	P45221 haemophilus
30	58.5	16.1	1044	1	BUB1_SCHPO	O94751 schizosacch
31	58	15.9	617	1	YACH_ECOLI	P36682 escherichia
32	58	15.9	874	1	SLAP_BACLI	P49052 bacillus li
33	57.5	15.8	334	1	G3P_BACST	P00362 bacillus st
34	57.5	15.8	335	1	G3P_BACCO	P15115 bacillus co
35	57.5	15.8	367	1	LHX4_MOUSE	P53776 mus musculu
36	57.5	15.8	419	1	Y223_MYCPN	P75465 mycoplasma
37	57.5	15.8	439	1	XYLA_LACLA	Q9cfg7 lactococcus
38	57.5	15.8	1852	1	CCAS_CYPCA	P22316 cyprinus ca
39	57	15.7	207	1	IL6_MARMO	O35736 marmota mon
40	57	15.7	256	1	YKJ9_YEAST	P34247 saccharomyc
41	57	15.7	355	1	CRTB_RHOSH	P54905 rhodobacter
42	56.5	15.5	470	1	RHBA_RHIME	Q9z3r2 rhizobium m
43	56.5	15.5	511	1	HUTH_VIBCH	Q9ksq4 vibrio chol
44	56.5	15.5	2190	1	CCAD_CHICK	O73700 gallus gall
45	56	15.4	336	1	Y625_METJA	Q58042 methanococc

ALIGNMENTS

RESULT 1
VAL1_TGMV
ID VAL1_TGMV STANDARD; PRT; 352 AA.
AC P03567;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein.
GN AC1.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10831;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;
RT "Complete nucleotide sequence of the infectious cloned DNA components
RT of tomato golden mosaic virus: potential coding regions and regulatory
RT sequences.";
RL EMBO J. 3:2197-2205(1984).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; K02029; -; NOT_ANNOTATED_CDS.
DR PIR; A04170; QQCVL1.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR ProDom; PD000736; Gemini_AL1; 1.
KW ATP-binding.
FT NP_BIND 223 230 ATP (POTENTIAL).
SQ SEQUENCE 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;

Query Match 100.0%; Score 364; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.8e-33;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
VAL1_PYMVV
ID VAL1_PYMVV STANDARD; PRT; 361 AA.

```
AC   P27258;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 protein.
OS   Potato yellow mosaic virus (isolate Venezuela).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10828;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91311403; PubMed=1856690;
RA   Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT   "The nucleotide sequence of the infectious cloned DNA components of
RT   potato yellow mosaic virus.";
RL   J. Gen. Virol. 72:1515-1520(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; D00940; BAA00782.1; -.
DR   PIR; JU0364; QQCVPT.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   361 AA;  40850 MW;  5627A33BF1264383 CRC64;
```

```
  Query Match             72.5%;  Score 264;  DB 1;  Length 361;
  Best Local Similarity   71.0%;  Pred. No. 2.9e-22;
  Matches   49;  Conservative   11;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: || ||:|||||||| || |||||||||: :|| |::||:||:|||:|||:|||: ||
Db    110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy     61 DRIFDKTPE 69
          |||| | ||
Db    170 DRIFMKAPE 178
```

```
RESULT   3
VAL1_CLVK
ID   VAL1_CLVK      STANDARD;      PRT;   358 AA.
AC   P14982;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (40.4 kDa protein).
GN   AC1.
OS   Cassava latent virus (strain West Kenyan 844).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10818;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Gay M.R.;
RT   "Nucleotide sequence of cassava latent virus DNA.";
RL   Nature 301:260-262(1983).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; J02057; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40346 MW;  ED173E753EE92D69 CRC64;
```

```
  Query Match             69.0%;  Score 251;  DB 1;  Length 358;
  Best Local Similarity   64.3%;  Pred. No. 8.3e-21;
  Matches   45;  Conservative   13;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||:|||||||| |::||| |:|||: || ||| :||| :|: :: |||||||||
Db    109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy     61 DRIFDKTPEP 70
          |||| : | |
Db    169 DRIFQEPPAP 178
```

```
RESULT   4
VAL1_CLVN
ID   VAL1_CLVN      STANDARD;      PRT;   358 AA.
AC   P14972;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (40.4 kDa protein).
GN   AC1.
OS   Cassava latent virus (strain Nigerian).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10819;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90174930; PubMed=2308831;
RA   Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT   "Nucleotide sequence of the infectious cloned DNA components of
RT   African cassava mosaic virus (Nigerian strain).";
RL   Nucleic Acids Res. 18:197-198(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X17095; CAA34953.1; -.
DR   PIR; S07594; S07594.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40435 MW;  1DB16BB0CB2D5E2C CRC64;
```

```
  Query Match             69.0%;  Score 251;  DB 1;  Length 358;
  Best Local Similarity   64.3%;  Pred. No. 8.3e-21;
  Matches   45;  Conservative   13;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||:|||||||| |::||| |:|||: || ||| :||| :|: :: |||||||||
Db    109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168
```

```
Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178
```

```
RESULT   5
VAL1_TYLCA
ID   VAL1_TYLCA     STANDARD;      PRT;   362 AA.
AC   P36279;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=36447;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93139778; PubMed=8423446;
RA   Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;
RT   "Nucleotide sequence and genome organization of tomato leaf curl
RT   geminivirus.";
RL   J. Gen. Virol. 74:147-151(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
DR   PIR; JQ1887; JQ1887.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   362 AA;  41197 MW;  343E7184B4704098 CRC64;
```

```
  Query Match             67.0%;  Score 244;  DB 1;  Length 362;
  Best Local Similarity   56.5%;  Pred. No. 5.1e-20;
  Matches   48;  Conservative    9;  Mismatches   12;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || |||||:|||||||| |::||| |:||| || ||| ::|| |: |: |||||||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194
```

```
RESULT   6
VAL1_TYLCU
ID   VAL1_TYLCU     STANDARD;      PRT;   359 AA.
AC   P38609;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=37139;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94256836; PubMed=8198442;
RA   Noris E., Hidalgo E., Accotto G., Moriones E.;
RT   "High similarity among the tomato yellow leaf curl virus isolates
RT   from the west Mediterranean basin: the nucleotide sequence of an
RT   infectious clone from Spain.";
RL   Arch. Virol. 135:165-170(1994).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z25751; CAA81026.1; -.
DR   PIR; S39211; S39211.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  41065 MW;  2D170A51EF80A3EC CRC64;
```

```
  Query Match             65.1%;  Score 237;  DB 1;  Length 359;
  Best Local Similarity   62.3%;  Pred. No. 3.1e-19;
  Matches   43;  Conservative    9;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || ||| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179
```

```
RESULT   7
VAL1_PHUV
ID   VAL1_PHUV      STANDARD;      PRT;   349 AA.
AC   Q06923;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AL1.
OS   Pepper huasteco virus (PHV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=28349;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94015007; PubMed=8409944;
RA   Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA   Rivera-Bustamante R.F.;
RT   "Complete nucleotide sequence of pepper huasteco virus: analysis and
RT   comparison with bipartite geminiviruses.";
RL   J. Gen. Virol. 74:2225-2231(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70418; CAA49856.1; -.
DR   PIR; S31875; S31875.
DR   PIR; JQ2300; JQ2300.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (BY SIMILARITY).
SQ   SEQUENCE   349 AA;  39722 MW;  D5F4E76CD56370F4 CRC64;
```

```
  Query Match             64.0%;  Score 233;  DB 1;  Length 349;
```

```
  Best Local Similarity   61.4%;  Pred. No. 8.3e-19;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| |::|| |:|||::| |||||||:|: |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   8
VAL1_TYLCM
ID   VAL1_TYLCM      STANDARD;      PRT;   359 AA.
AC   P27260;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10833;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92107660; PubMed=1840676;
RA   Kheyr-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,
RA   Gronenborn B.;
RT   "Tomato yellow leaf curl virus from Sardinia is a
RT   whitefly-transmitted monopartite geminivirus.";
RL   Nucleic Acids Res. 19:6763-6769(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X61153; CAA43466.1; -.
DR   PIR; S22593; S22593.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  40733 MW;  9717B4A07C93EFA7 CRC64;


  Query Match             63.2%;  Score 230;  DB 1;  Length 359;
  Best Local Similarity   59.4%;  Pred. No. 1.9e-18;
  Matches   41;  Conservative   11;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   9
VAL1_BGMV
ID   VAL1_BGMV      STANDARD;      PRT;   353 AA.
AC   P05175;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   AL1 protein (40.2 kDa protein).
GN   AC1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Howarth A.J., Caton J., Bossert M., Goodman R.M.;
RT   "Nucleotide sequence of bean golden mosaic virus and a model for gene
RT   regulation in geminiviruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M10070; AAA46318.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   353 AA;  40190 MW;  80FA779DF6029A34 CRC64;


  Query Match             62.6%;  Score 228;  DB 1;  Length 353;
  Best Local Similarity   61.4%;  Pred. No. 3.1e-18;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:|||||||||| |::||: |:|||| | | || |::|: |: |: | ||: |||
Db     110 TIEWGQFQVDGRSARGGQQSANDSYAKALNADSIESALTILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFFKVPEP 179


RESULT  10
VAL1_ABMVW
ID   VAL1_ABMVW      STANDARD;      PRT;   355 AA.
AC   P21947;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AC1.
OS   Abutilon mosaic virus (isolate West India).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10816;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91020984; PubMed=2219703;
RA   Frischmuth T., Zimmat G., Jeske H.;
RT   "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT   as well as eukaryotic features.";
RL   Virology 178:461-468(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
```

```
DR   EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR   PIR; A36214; QQCVW1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   355 AA;  40257 MW;  16A2CA8A63251E95 CRC64;


  Query Match             62.4%;  Score 227;  DB 1;  Length 355;
  Best Local Similarity   60.0%;  Pred. No. 4e-18;
  Matches   42;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |  |||||:|||||||| ||:||: |:|||   : || |::|: |: |: | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
VAL1_BCTV
ID   VAL1_BCTV      STANDARD;      PRT;   358 AA.
AC   P14991;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 protein (40.8 kDa protein).
OS   Beet curly top virus (BCTV).
OC   Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
OX   NCBI_TaxID=10840;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT   "The nucleotide sequence of an infectious clone of the geminivirus
RT   beet curly top virus.";
RL   EMBO J. 5:1761-1767(1986).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40889 MW;  39A45FE3C0B9C333 CRC64;


  Query Match             61.5%;  Score 224;  DB 1;  Length 358;
  Best Local Similarity   57.1%;  Pred. No. 8.7e-18;
  Matches   40;  Conservative   16;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| ||:||: |:|||:| ::||||::|: |: |  | ||| :|
Db     110 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 169

Qy      61 DRIFDKTPEP 70
           :|| : |:|
Db     170 QKIFQRPPDP 179


RESULT  12
VAL1_TMOV
ID   VAL1_TMOV      STANDARD;      PRT;   361 AA.
AC   Q06657;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AL1.
OS   Tomato mottle virus (isolate Florida) (TMoV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=36449;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93107858; PubMed=1469361;
RA   Abouzid A.M., Polston J.E., Hiebert E.;
RT   "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT   isolated from tomatoes in Florida.";
RL   J. Gen. Virol. 73:3225-3229(1992).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L14460; AAC32414.1; -.
DR   PIR; JQ1870; JQ1870.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (BY SIMILARITY).
SQ   SEQUENCE   361 AA;  40516 MW;  8138B65CEEAC6950 CRC64;


  Query Match             61.5%;  Score 224;  DB 1;  Length 361;
  Best Local Similarity   57.1%;  Pred. No. 8.8e-18;
  Matches   40;  Conservative   16;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||: |:|||| | : || ::||: |: :: | ||: |||
Db     110 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  13
VAL1_TYLCV
ID   VAL1_TYLCV     STANDARD;      PRT;   357 AA.
AC   P27259;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10832;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92024070; PubMed=1926771;
RA   Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;
RT   "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RT   with a single genomic component.";
RL   Virology 185:151-161(1991).
```

```
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15656; CAA33688.1; -.
DR   PIR; D40779; QQCVC1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     219    226       ATP (POTENTIAL).
SQ   SEQUENCE   357 AA;  40678 MW;  939AB68E1AB3B2A7 CRC64;


  Query Match             59.1%;  Score 215;  DB 1;  Length 357;
  Best Local Similarity   67.2%;  Pred. No. 8.8e-17;
  Matches   41;  Conservative   10;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           :|  |:|||||||| |::||| |||||: || ||| |::|| |: |: ||||||:|||||
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171


RESULT  14
VAL1_SLCV
ID   VAL1_SLCV      STANDARD;      PRT;   347 AA.
AC   P29048;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   AL1 protein.
OS   Squash leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10829;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91082449; PubMed=1984668;
RA   Lazarowitz S.G., Lazdins I.B.;
RT   "Infectivity and complete nucleotide sequence of the cloned genomic
RT   components of a bipartite squash leaf curl geminivirus with a broad
RT   host range phenotype.";
RL   Virology 180:58-69(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M38183; AAC32410.1; ALT_INIT.
DR   PIR; C36785; QQCVS1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     218    225       ATP (POTENTIAL).
SQ   SEQUENCE   347 AA;  39110 MW;  AFDAEBDDE122110E CRC64;


  Query Match             34.1%;  Score 124;  DB 1;  Length 347;
  Best Local Similarity   37.9%;  Pred. No. 1.3e-06;
  Matches   25;  Conservative   14;  Mismatches   23;  Indels    4;  Gaps    1;

Qy       5 GEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRIF 64
           |:::| |    |     :|    |:|| |  ||| ||:   |: ::  :||| :|::|:|
Db     116 GQYKVSG----GSKSNKDDVYHNAVNAGSAGEALDIIKAGDPKTFIVNYHNLLANVERLF 171

Qy      65 DKTPEP 70
             | ||| .
Db     172 QKPPEP 177


RESULT  15
Y175_HELPJ
ID   Y175_HELPJ     STANDARD;      PRT;   299 AA.
AC   Q9ZMQ7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein JHP0161 precursor.
GN   JHP0161.
OS   Helicobacter pylori J99 (Campylobacter pylori J99).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
OX   NCBI_TaxID=85963;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99120557; PubMed=9923682;
RA   Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA   Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA   Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA   Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA   Trust T.J.;
RT   "Genomic sequence comparison of two unrelated isolates of the human
RT   gastric pathogen Helicobacter pylori.";
RL   Nature 397:176-180(1999).
CC   -!- SIMILARITY: BELONGS TO THE PPIC/PARVULIN FAMILY OF ROTAMASES.
CC       STRONG, TO C.JEJUNI CBF2.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001454; AAD05744.1; -.
DR   HSSP; Q13526; 1PIN.
DR   InterPro; IPR000297; Rotamase.
DR   Pfam; PF00639; Rotamase; 1.
DR   PROSITE; PS01096; PPIC_PPIASE_1; 1.
DR   PROSITE; PS50198; PPIC_PPIASE_2; 1.
KW   Hypothetical protein; Isomerase; Rotamase; Signal; Complete proteome.
FT   SIGNAL       1     21       POTENTIAL.
FT   CHAIN       22    299       HYPOTHETICAL PROTEIN JHP0161.
FT   DOMAIN     154    253       PPIC.
SQ   SEQUENCE   299 AA;  34040 MW;  9C037B1CD1110143 CRC64;


  Query Match             17.7%;  Score 64.5;  DB 1;  Length 299;
  Best Local Similarity   28.1%;  Pred. No. 5.1;
  Matches   18;  Conservative   14;  Mismatches   11;  Indels   21;  Gaps    2;

Qy      22 NDAAAEALNAS----------------SKEEALQIIREKIPEKYLFQFHNLNSNLDR 62
           |:| || || :                 :|::| :: : :|||| :  |:  |:| |:
Db      93 NEAKAEKLNQTPEFKAMMEAVKKQALVEFWAKKQAEEVKKIQIPEKEMQDFY--NANKDQ 150

Qy      63 IFDK 66
           :| |
Db     151 LFVK 154
```

Search completed: October 10, 2002, 02:09:49
Job time: 1347 sec

wt = DK
seq

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:39:22 ; Search time 69.63 Seconds
(without alignments)
96.600 Million cell updates/sec

Title: US-09-289-346A-1
Perfect score: 364
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	364	100.0	352	1	QQCVL1	AL1 protein - toma
2	264	72.5	361	1	QQCVPT	AL1 protein - pota
3	251	69.0	358	2	S07594	hypothetical prote
4	244	67.0	362	1	JQ1887	AL1 protein - toma
5	237	65.1	359	2	S39211	gene C1 protein -
6	233	64.0	349	2	JQ2300	replicase - pepper
7	233	64.0	349	2	S31875	AL1 protein - pepp
8	230	63.2	359	2	S22593	hypothetical prote
9	227	62.4	351	2	JQ2327	AL1 protein - Indi
10	227	62.4	355	1	QQCVW1	AV1 protein - abut
11	224	61.5	358	1	JQ1870	AL1 protein - toma
12	224	61.5	359	2	S39235	gene C1 protein -
13	224	61.5	385	2	S28360	AL1 protein - beet
14	217	59.6	360	2	S59885	replication-associ
15	215	59.1	357	1	QQCVC1	AL1 protein - toma
16	142	39.0	131	2	S45059	AC1 protein (clone
17	124	34.1	347	1	QQCVS1	AL1 protein - squa
18	68	18.7	587	2	C95382	Fc gamma (IgG) rec
19	66	18.1	714	2	B71967	probable ferrichro
20	64.5	17.7	299	2	A55347	probable peptidyl-
21	64.5	17.7	1713	2	A46227	adhesive ligand ep
22	63.5	17.4	1610	2	JH0422	voltage-dependent
23	63.5	17.4	1646	2	JH0564	voltage-dependent
24	63.5	17.4	2161	2	A38198	calcium channel al
25	63.5	17.4	2181	2	T42742	calcium channel al
26	63.5	17.4	2203	2	T12544	voltage-dependent
27	62.5	17.2	447	2	D90558	hypothetical prote
28	62.5	17.2	642	2	AD3475	topoisomerase iv s
29	62	17.0	154	2	G97976	nitrogen regulator
30	62	17.0	181	2	D42452	conserved hypothet
31	62	17.0	295	2	A70091	C1 protein - tobac
32	61.5	16.9	481	2	T36175	probable phosphoes
33	61	16.8	840	2	RGBYH1	probable large ATP
34	61	16.8	1502	1	A75087	CYC1/CYP3 transcri
35	60.5	16.6	354	2	T48959	acetyl ornithine d
36	60.5	16.6	1229	2	S52437	kinesin-like prote
37	60	16.5	447	2	E97700	CDP-diacylglycerol
38	59.5	16.3	1033	2	T05244	hypothetical prote
39	59.5	16.3	2137	2	G82060	hypothetical prote
40	59	16.2	160	2	C82085	hypothetical prote
41	59	16.2	316	2	AG3582	conserved hypothet
42	59	16.2	338	2	B97011	iron(III)-binding
43	59	16.2	340	2	B71078	probably O-sialogl
44	59	16.2	397	2	A82892	probable NADH oxid
45	59	16.2	416	2		hypothetical prote

ALIGNMENTS

RESULT 1
QQCVL1
AL1 protein - tomato golden mosaic virus
C;Species: tomato golden mosaic virus
A;Note: host Nicotiana sp. (tobacco)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C;Accession: A04170
R;Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.
EMBO J. 3, 2197-2205, 1984
A;Title: Complete nucleotide sequence of the infectious cloned DNA components of toma
A;Reference number: A04163
A;Accession: A04170
A;Molecule type: DNA
A;Residues: 1-352 <HAM>
C;Comment: The genome consists of two circular, single-stranded DNA components, DNA A
.
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 100.0%; Score 364; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 3.1e-33;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
QQCVPT
AL1 protein - potato yellow mosaic virus (isolate Venezuela)
C;Species: potato yellow mosaic virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C;Accession: JU0364
R;Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.
J. Gen. Virol. 72, 1515-1520, 1991
A;Title: The nucleotide sequence of the infectious cloned DNA components of potato ye
A;Reference number: JU0362; MUID:91311403
A;Accession: JU0364
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-361 <COU>
A;Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             72.5%;  Score 264;  DB 1;  Length 361;
  Best Local Similarity   71.0%;  Pred. No. 5.7e-22;
  Matches   49;  Conservative   11;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:|||||||| || |||||||||: :|| |::||:||:|||:|||:|||: ||
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 DRIFDKTPE 69
           |||| | ||
Db     170 DRIFMKAPE 178


RESULT   3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C;Species: cassava latent virus
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C;Accession: S07594
R;Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A;Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A;Reference number: S07590; MUID:90174930
A;Accession: S07594
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <MOR>
A;Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C;Genetics:
A;Map position: segment DNA1
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             69.0%;  Score 251;  DB 2;  Length 358;
  Best Local Similarity   64.3%;  Pred. No. 1.6e-20;
  Matches   45;  Conservative   13;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||| |:|||: || ||| :||| :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   4
JQ1887
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1887
R;Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A;Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A;Reference number: JQ1885; MUID:93139778
A;Accession: JQ1887
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-362 <DRY>
A;Cross-references: GB:S53251
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             67.0%;  Score 244;  DB 1;  Length 362;
  Best Local Similarity   56.5%;  Pred. No. 1e-19;
  Matches   48;  Conservative    9;  Mismatches   12;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || |||||:|||||||| |::||| |:|||  || ||| ::||  |: |: |||||||||

Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                  |||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   5
S39211
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C;Accession: S39211
R;Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A;Description: High similarity among the tomato yellow leaf curl virus isolates from
A;Reference number: S39209
A;Accession: S39211
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <NOR>
A;Cross-references: EMBL:Z25751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             65.1%;  Score 237;  DB 2;  Length 359;
  Best Local Similarity   62.3%;  Pred. No. 6.2e-19;
  Matches   43;  Conservative    9;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || ||| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   6
JQ2300
replicase - pepper huasteco virus (component A)
N;Alternate names: ORF AL1 protein
C;Species: pepper huasteco virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C;Accession: JQ2300
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A;Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A;Reference number: JQ2299; MUID:94015007
A;Accession: JQ2300
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             64.0%;  Score 233;  DB 2;  Length 349;
  Best Local Similarity   61.4%;  Pred. No. 1.7e-18;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||||||:|||||||| |::||  |:|||::| |||||||:|: |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   7
S31875
AL1 protein - pepper rizado amarillo virus
C;Species: pepper rizado amarillo virus
```

C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Sep-1999
C;Accession: S31875
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera-Bu
submitted to the EMBL Data Library, February 1993
A;Description: Complete nucleotide sequence of pepper huasteco virus: analysis and compa
A;Reference number: S31872
A;Accession: S31875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
A;Note: the source is designated as pepper huasteco virus
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 64.0%; Score 233; DB 2; Length 349;
Best Local Similarity 61.4%; Pred. No. 1.7e-18;
Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

Qy 1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
Db 110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy 61 DRIFDKTPEP 70
Db 170 NRIFQTPPEP 179

RESULT 8
S22593
hypothetical protein C4 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Sep-1999
C;Accession: S22593
R;Kheyr-Pour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.
Nucleic Acids Res. 19, 6763-6769, 1991
A;Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monoparti
A;Reference number: S22588; MUID:92107660
A;Accession: S22593
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-359 <KHE>
A;Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 63.2%; Score 230; DB 2; Length 359;
Best Local Similarity 59.4%; Pred. No. 3.8e-18;
Matches 41; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

Qy 2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
Db 111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy 62 RIFDKTPEP 70
Db 171 KVFQVPPAP 179

RESULT 9
JQ2327
AL1 protein - Indian cassava mosaic virus
N;Alternate names: replication-associated protein
C;Species: Indian cassava mosaic virus
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C;Accession: JQ2327; S35883
R;Hong, Y.G.; Robinson, D.J.; Harrison, B.D.
J. Gen. Virol. 74, 2437-2443, 1993
A;Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-tran
A;Reference number: JQ2326; MUID:94065670
A;Accession: JQ2327
A;Molecule type: DNA
A;Residues: 1-351 <HON>
A;Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 62.4%; Score 227; DB 2; Length 351;
Best Local Similarity 64.2%; Pred. No. 8.1e-18;
Matches 43; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

Qy 4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
Db 113 WGTFQIDGRSARGGQQSANDAYAAALNSGSKSEALKILRELAPRDYLRDFHHISSNLDRI 172

Qy 64 FDKTPEP 70
Db 173 FTKPPPP 179

RESULT 10
QQCVW1
AV1 protein - abutilon mosaic virus (isolate West India)
C;Species: abutilon mosaic virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Apr-1994
C;Accession: A36214
R;Frischmuth, T.; Zimmat, G.; Jeske, H.
Virology 178, 461-468, 1990
A;Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as
A;Reference number: A36214; MUID:91020984
A;Accession: A36214
A;Molecule type: DNA
A;Residues: 1-355 <FRI>
A;Cross-references: EMBL:X15983
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 62.4%; Score 227; DB 1; Length 355;
Best Local Similarity 60.0%; Pred. No. 8.2e-18;
Matches 42; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

Qy 1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
Db 110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy 61 DRIFDKTPEP 70
Db 170 ERIFAKAPEP 179

RESULT 11
JQ1870
AL1 protein - tomato mottle virus (isolate Florida)
C;Species: tomato mottle virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1870
R;Abouzid, A.M.; Polston, J.E.; Hiebert, E.
J. Gen. Virol. 73, 3225-3229, 1992
A;Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated f
A;Reference number: JQ1869; MUID:93107858
A;Accession: JQ1870
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <ABO>
A;Cross-references: GB:L14460
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 61.5%; Score 224; DB 1; Length 358;
Best Local Similarity 57.1%; Pred. No. 1.8e-17;
Matches 40; Conservative 16; Mismatches 14; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||: |:|||| | : || ::||: |: :: | ||: |||
Db     107 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 166

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     167 ERIFAKAPEP 176
```

RESULT 12
S39235
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Sep-1997
C;Accession: S39235
R;Crespi, S.; Noris, E.; Vaira, A.; Bosco, D.; Accotto, G.
submitted to the EMBL Data Library, December 1993
A;Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.
A;Reference number: S39233
A;Accession: S39235
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <CRE>
A;Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             61.5%;  Score 224;  DB 2;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 1.8e-17;
  Matches   40;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || ||| :|:: |  |:  |||::||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINARSKSEALDVIKQLAPRDYVLHFHNISSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179
```

RESULT 13
S28360
AL1 protein - beet curly top virus
C;Species: beet curly top virus
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
C;Accession: S28360
R;Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
EMBO J. 5, 1761-1767, 1986
A;Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly to
A;Reference number: S28360
A;Accession: S28360
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-385 <STA>
A;Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             61.5%;  Score 224;  DB 2;  Length 385;
  Best Local Similarity   57.1%;  Pred. No. 1.9e-17;
  Matches   40;  Conservative   16;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| ||:||: |:|||||:| ::||||::|: |: |  | ||| :|
Db     137 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 196

Qy      61 DRIFDKTPEP 70
           :|| : |:|
Db     197 QKIFQRPPDP 206
```

RESULT 14
S59885
replication-associated protein C1 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
C;Accession: S59885
R;Hong, Y.; Harrison, B.D.
submitted to the EMBL Data Library, February 1995
A;Description: Nucleotide sequences from tomato leaf curl viruses from different coun
d geminiviruses.
A;Reference number: S58346
A;Accession: S59885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <HON>
A;Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA88229.1; PID:g974211
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             59.6%;  Score 217;  DB 2;  Length 360;
  Best Local Similarity   62.1%;  Pred. No. 1.1e-16;
  Matches   41;  Conservative   12;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           :|' ||:|||||||| |::||| |||:|: ||  || |:||| |: :: |||||:|||||
Db     113 FGVFQIDGRSARGGQQSANDAYAEAINSGSKAAALDILREKAPKDFVLQFHNLNANLDRI 172

Qy      64 FDKTPE 69
           |  : |
Db     173 FTPSAE 178
```

RESULT 15
QQCVC1
AL1 protein - tomato yellow leaf curl virus
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C;Accession: D40779
R;Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.
Virology 185, 151-161, 1991
A;Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin
A;Reference number: A40779; MUID:92024070
A;Accession: D40779
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-357 <NAV>
A;Cross-references: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             59.1%;  Score 215;  DB 1;  Length 357;
  Best Local Similarity   67.2%;  Pred. No. 1.8e-16;
  Matches   41;  Conservative   10;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           :|  |:|||||||| |::||| ||||| : || ||| |::|| |: |: ||||||||||
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171
```

Search completed: October 10, 2002, 01:48:42
Job time: 560 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:44:57 ; Search time 407.82 Seconds
(without alignments)
47.487 Million cell updates/sec

Title: US-09-289-346A-1
Perfect score: 364
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1006125 seqs, 276659714 residues

Total number of hits satisfying chosen parameters: 1006125

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	74	20.3	340	6	US-10-219-999-59588	Sequence 59588, A
2	71.5	19.6	945	5	US-09-791-537-132140	Sequence 132140,
3	68	18.7	587	5	US-09-791-537-33198	Sequence 33198, A
4	67.5	18.5	512	5	US-09-791-537-52432	Sequence 52432, A
5	67.5	18.5	1449	5	US-09-791-537-52442	Sequence 52442, A
6	65	17.9	293	6	US-10-151-336-3	Sequence 3, Appli
7	65	17.9	335	6	US-10-151-336-6	Sequence 6, Appli
8	64.5	17.7	299	5	US-09-791-537-79519	Sequence 79519, A
9	64.5	17.7	1238	5	US-09-791-537-144585	Sequence 144585,
10	64.5	17.7	1331	5	US-09-791-537-118290	Sequence 118290,
11	64.5	17.7	1713	5	US-09-643-992A-2	Sequence 2, Appli
12	64.5	17.7	1713	5	US-09-791-537-87556	Sequence 87556, A
13	64.5	17.7	1713	6	US-10-171-311-113	Sequence 113, App
14	64.5	17.7	1806	5	US-09-791-537-53064	Sequence 53064, A
15	64.5	17.7	3298	7	US-60-365-384-181	Sequence 181, App
16	64.5	17.7	3332	7	US-60-365-384-182	Sequence 182, App
17	63.5	17.4	2181	1	PCT-US02-23913-50	Sequence 50, Appl
18	63.5	17.4	2181	1	PCT-US01-32045-288	Sequence 288, App
19	63.5	17.4	2181	6	US-10-205-823-50	Sequence 50, Appl
20	62.5	17.2	447	5	US-09-991-150-369	Sequence 369, App
21	62.5	17.2	447	5	US-09-989-328-369	Sequence 369, App
22	62.5	17.2	447	6	US-10-103-295-129	Sequence 129, App
23	62.5	17.2	447	6	US-10-121-062-254	Sequence 254, App
24	62.5	17.2	447	6	US-10-173-689-254	Sequence 254, App
25	62.5	17.2	447	6	US-10-173-690-254	Sequence 254, App
26	62.5	17.2	447	6	US-10-173-691-254	Sequence 254, App
27	62.5	17.2	447	6	US-10-173-692-254	Sequence 254, App
28	62.5	17.2	447	6	US-10-173-695-254	Sequence 254, App
29	62.5	17.2	447	6	US-10-173-697-254	Sequence 254, App
30	62.5	17.2	447	6	US-10-173-698-254	Sequence 254, App
31	62.5	17.2	447	6	US-10-173-699-254	Sequence 254, App
32	62.5	17.2	447	6	US-10-173-701-254	Sequence 254, App
33	62.5	17.2	447	6	US-10-173-704-254	Sequence 254, App
34	62.5	17.2	447	6	US-10-173-708-254	Sequence 254, App
35	62.5	17.2	447	6	US-10-174-571-254	Sequence 254, App
36	62.5	17.2	447	6	US-10-174-574-254	Sequence 254, App
37	62.5	17.2	447	6	US-10-174-583-254	Sequence 254, App
38	62.5	17.2	447	6	US-10-173-694-254	Sequence 254, App
39	62.5	17.2	447	6	US-10-173-707-254	Sequence 254, App
40	62.5	17.2	447	6	US-10-174-569-254	Sequence 254, App
41	62.5	17.2	447	6	US-10-174-570-254	Sequence 254, App
42	62.5	17.2	447	6	US-10-174-578-254	Sequence 254, App
43	62.5	17.2	447	6	US-10-174-582-254	Sequence 254, App
44	62.5	17.2	447	6	US-10-174-585-254	Sequence 254, App
45	62.5	17.2	447	6	US-10-174-586-254	Sequence 254, App

ALIGNMENTS

RESULT 1
US-10-219-999-59588
; Sequence 59588, Application US/10219999
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: cDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)C
; CURRENT APPLICATION NUMBER: US/10/219,999
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/324,109
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/312,544
; PRIOR FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 63520
; SEQ ID NO 59588
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Zea mays
US-10-219-999-59588

Query Match 20.3%; Score 74; DB 6; Length 340;
Best Local Similarity 26.8%; Pred. No. 1.2;
Matches 19; Conservative 15; Mismatches 21; Indels 16; Gaps 3;

```
Qy      15 RGGCQTSNDAAAEAL-NASSKEEALQ--------IIREKIPEKYLFQFHN-------LNS 58
           | || :    |  |  : ::|::  |            :::|::|: : |  ||       |||
Db      30 RPGCSSRTPTATSATHHGATKQQVRQGKVSDETKMVQEELPQAWWFDSHNLARPSPWLNS 89

Qy      59 NLDRIFDKTPE 69
             |  : ||| :
Db      90 TLSELDDKTKQ 100
```

RESULT 2
US-09-791-537-132140
; Sequence 132140, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210

```
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 132140
;   LENGTH: 945
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-791-537-132140


  Query Match             19.6%;  Score 71.5;  DB 5;  Length 945;
  Best Local Similarity   34.4%;  Pred. No. 9;
  Matches   22;  Conservative    8;  Mismatches   19;  Indels   15;  Gaps    2;

Qy      10 DGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIP-------------EKYLFQFHN 55
           :|    || |::|  || ||   ||||  :| |||               | :|  |:|
Db     145 NGAGVAGGYQSNNYNAA-ALGMLSKEERAEIQREKAKNPGRNLVKPKWENLEPFLKDFYN 203

Qy      56 LNSN 59
           :: |
Db     204 IHPN 207


RESULT   3
US-09-791-537-33198
; Sequence 33198, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 33198
;   LENGTH: 587
;   TYPE: PRT
;   ORGANISM: Streptococcus sp
US-09-791-537-33198


  Query Match             18.7%;  Score 68;  DB 5;  Length 587;
  Best Local Similarity   30.0%;  Pred. No. 13;
  Matches   18;  Conservative   13;  Mismatches   21;  Indels    8;  Gaps    1;

Qy      12 RSARGGCQTSNDAAAEALNASSKE--------EALQIIREKIPEKYLFQFHNLNSNLDRI 63
           :|| || :|: |  |: || :|::         :|||   |   :| |   :| |  : ::
Db     114 KSANGGLKTALDDMADTLNQASRKVNELSAQNKALQAEAEAAAQKALDALNNKNEQIAKL 173


RESULT   4
US-09-791-537-52432
; Sequence 52432, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 52432
;   LENGTH: 512
;   TYPE: PRT
;   ORGANISM: Human immunodeficiency virus type 1
US-09-791-537-52432


  Query Match             18.5%;  Score 67.5;  DB 5;  Length 512;
  Best Local Similarity   27.4%;  Pred. No. 13;
  Matches   23;  Conservative   11;  Mismatches   33;  Indels   17;  Gaps    3;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKY----------L 50
           |||   :::  :   |   : |  |: ||    | ||:   |::   |
Db      34 LVWASRELERFACNPGLMDTADGCAKLLN--QLEPALKTGSEELRSLYNALAVLYCVHSR 91

Qy      51 FQFHNLNSNLDRIFDK----TPEP 70
           | ||    ||:| :|     |||
Db      92 IQIHNTQEALDKIKEKQEQHKPEP 115


RESULT   5
US-09-791-537-52442
; Sequence 52442, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 52442
;   LENGTH: 1449
;   TYPE: PRT
;   ORGANISM: Human immunodeficiency virus type 1
US-09-791-537-52442


  Query Match             18.5%;  Score 67.5;  DB 5;  Length 1449;
  Best Local Similarity   27.4%;  Pred. No. 48;
  Matches   23;  Conservative   11;  Mismatches   33;  Indels   17;  Gaps    3;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKY----------L 50
           |||   :::  :   |   : |  |: ||    | ||:   |::   |
Db      34 LVWASRELERFACNPGLMDTADGCAKLLN--QLEPALKTGSEELRSLYNALAVLYCVHSR 91

Qy      51 FQFHNLNSNLDRIFDK----TPEP 70
           | ||    ||:| :|     |||
Db      92 IQIHNTQEALDKIKEKQEQHKPEP 115


RESULT   6
US-10-151-336-3
; Sequence 3, Application US/10151336
; GENERAL INFORMATION:
;  APPLICANT: Mason, Hugh
;  APPLICANT:  Palmer, Kenneth
;  APPLICANT:  Hefferon, Kathleen
;  APPLICANT:  Mor, Tsafrir
;  APPLICANT:  Arntzen, Charles
;  TITLE OF INVENTION: Gemini Virus Vectors for Gene Expression in Plants
;  FILE REFERENCE: 4868/84453
;  CURRENT APPLICATION NUMBER: US/10/151,336
;  CURRENT FILING DATE:  2002-05-20
;  PRIOR APPLICATION NUMBER: US/09/414,276
;  PRIOR FILING DATE: 1999-10-07
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 293
;   TYPE: PRT
;   ORGANISM: bean yellow dwarf virus
US-10-151-336-3
```

```
  Query Match             17.9%;  Score 65;  DB 6;  Length 293;
  Best Local Similarity   33.3%;  Pred. No. 13;
  Matches   13;  Conservative   11;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      31 ASSKEEALQIIREKIPEKYLFQFHNLNSNLDRIFDKTPE 69
           |:|||| | :|:|: | ::  :   |  : :::|   ||
Db     133 ATSKEEYLDMIKEEFPHEWATKLQWLEYSANKLFPPQPE 171


RESULT   7
US-10-151-336-6
; Sequence 6, Application US/10151336
; GENERAL INFORMATION:
;  APPLICANT: Mason, Hugh
;  APPLICANT:  Palmer, Kenneth
;  APPLICANT:  Hefferon, Kathleen
;  APPLICANT:  Mor, Tsafrir
;  APPLICANT:  Arntzen, Charles
;  TITLE OF INVENTION: Gemini Virus Vectors for Gene Expression in Plants
;  FILE REFERENCE: 4868/84453
;  CURRENT APPLICATION NUMBER: US/10/151,336
;  CURRENT FILING DATE:  2002-05-20
;  PRIOR APPLICATION NUMBER: US/09/414,276
;  PRIOR FILING DATE: 1999-10-07
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 335
;   TYPE: PRT
;   ORGANISM: bean yellow dwarf virus
US-10-151-336-6


  Query Match             17.9%;  Score 65;  DB 6;  Length 335;
  Best Local Similarity   33.3%;  Pred. No. 15;
  Matches   13;  Conservative   11;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      31 ASSKEEALQIIREKIPEKYLFQFHNLNSNLDRIFDKTPE 69
           |:|||| | :|:|: | ::  :   |  : :::|   ||
Db     133 ATSKEEYLDMIKEEFPHEWATKLQWLEYSANKLFPPQPE 171


RESULT   8
US-09-791-537-79519
; Sequence 79519, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 79519
;   LENGTH: 299
;   TYPE: PRT
;   ORGANISM: Helicobacter pylori
US-09-791-537-79519


  Query Match             17.7%;  Score 64.5;  DB 5;  Length 299;
  Best Local Similarity   28.1%;  Pred. No. 15;
  Matches   18;  Conservative   14;  Mismatches   11;  Indels   21;  Gaps    2;

Qy      22 NDAAAEALNAS----------------SKEEALQIIREKIPEKYLFQFHNLNSNLDR 62
           |:| || || :                :|::| :: : :|||| :  |:  |:| |:
Db      93 NEAKAEKLNQTPEFKAMMEAVKKQALVEFWAKKQAEEVKKIQIPEKEMQDFY--NANKDQ 150

Qy      63 IFDK 66
           :| |
Db     151 LFVK 154


RESULT   9
US-09-791-537-144585
; Sequence 144585, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 144585
;   LENGTH: 1238
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-791-537-144585


  Query Match             17.7%;  Score 64.5;  DB 5;  Length 1238;
  Best Local Similarity   34.4%;  Pred. No. 93;
  Matches   21;  Conservative   12;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPEKYLFQFHNLNSNLDRIFD 65
           ||  :|      :  || :| |  ||: | ||| :|:| :| |       |:|| |:: :
Db     553 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLLN 608

Qy      66 K 66
           :
Db     609 E 609


RESULT  10
US-09-791-537-118290
; Sequence 118290, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 118290
;   LENGTH: 1331
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-791-537-118290


  Query Match             17.7%;  Score 64.5;  DB 5;  Length 1331;
  Best Local Similarity   34.4%;  Pred. No. 1e+02;
  Matches   21;  Conservative   12;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPEKYLFQFHNLNSNLDRIFD 65
           ||  :|      :  || :| |  ||: | ||| :|:| :| |       |:|| |:: :
Db     646 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLLN 701

Qy      66 K 66
           :
Db     702 E 702


RESULT  11
US-09-643-992A-2
```

```
; Sequence 2, Application US/09643992A
; GENERAL INFORMATION:
;  APPLICANT: Oriental Yeast Co., Ltd.
;  APPLICANT:  NAITOU Toshikuni
;  TITLE OF INVENTION: Method for neuranagenesis
;  FILE REFERENCE: FP/O-21-8US
;  CURRENT APPLICATION NUMBER: US/09/643,992A
;  CURRENT FILING DATE:  2000-08-23
;  NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 2
;   LENGTH: 1713
;   TYPE: PRT
;   ORGANISM: Human
US-09-643-992A-2


  Query Match             17.7%;  Score 64.5;  DB 5;  Length 1713;
  Best Local Similarity   34.4%;  Pred. No. 1.4e+02;
  Matches   21;  Conservative   12;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPEKYLFQFHNLNSNLDRIFD 65
           ||  :|      :  || :| |  ||: | ||| :|:| :| |       |:|| |:: :
Db     553 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLLN 608

Qy      66 K 66
           :
Db     609 E 609


RESULT  12
US-09-791-537-87556
; Sequence 87556, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 87556
;   LENGTH: 1713
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-791-537-87556


  Query Match             17.7%;  Score 64.5;  DB 5;  Length 1713;
  Best Local Similarity   34.4%;  Pred. No. 1.4e+02;
  Matches   21;  Conservative   12;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPEKYLFQFHNLNSNLDRIFD 65
           ||  :|      :  || :| |  ||: | ||| :|:| :| |       |:|| |:: :
Db     553 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLLN 608

Qy      66 K 66
           :
Db     609 E 609


RESULT  13
US-10-171-311-113
; Sequence 113, Application US/10171311
; GENERAL INFORMATION:
;  APPLICANT: Schlegel, Robert
;  APPLICANT:  Chen, Yan
;  APPLICANT:  Zhao, Xumei
;  APPLICANT:  Monahan, John
;  APPLICANT:  Kamatkar, Shubhangi
;  APPLICANT:  Glatt, Karen
;  APPLICANT:  Gannavarapu, Manjula
;  APPLICANT:  Hoersh, Sebastian
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;  TITLE OF INVENTION:  IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
;  TITLE OF INVENTION:  OF CERVICAL CANCER
;  FILE REFERENCE: MRI-035
;  CURRENT APPLICATION NUMBER: US/10/171,311
;  CURRENT FILING DATE:  2002-06-12
;  PRIOR APPLICATION NUMBER: US 60/298,159
;  PRIOR FILING DATE: 2001-06-13
;  PRIOR APPLICATION NUMBER: US 60/298,155
;  PRIOR FILING DATE: 2001-06-13
;  PRIOR APPLICATION NUMBER: US 60/335,936
;  PRIOR FILING DATE: 2001-11-14
;  NUMBER OF SEQ ID NOS: 238
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 113
;   LENGTH: 1713
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-171-311-113


  Query Match             17.7%;  Score 64.5;  DB 6;  Length 1713;
  Best Local Similarity   34.4%;  Pred. No. 1.4e+02;
  Matches   21;  Conservative   12;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPEKYLFQFHNLNSNLDRIFD 65
           ||  :|      :  || :| |  ||: | ||| :|:| :| |       |:|| |:: :
Db     553 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLLN 608

Qy      66 K 66
           :
Db     609 E 609


RESULT  14
US-09-791-537-53064
; Sequence 53064, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 53064
;   LENGTH: 1806
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-791-537-53064


  Query Match             17.7%;  Score 64.5;  DB 5;  Length 1806;
  Best Local Similarity   34.4%;  Pred. No. 1.5e+02;
  Matches   21;  Conservative   12;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPEKYLFQFHNLNSNLDRIFD 65
           ||  :|      :  || :| |  ||: | ||| :|:| :| |       |:|| |:: :
Db     646 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLLN 701

Qy      66 K 66
           :
Db     702 E 702


RESULT  15
US-60-365-384-181
; Sequence 181, Application US/60365384
```

; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Wang, Jian-rui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ghosh, Malabika
; APPLICANT: Zhao, Qing A.
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Weng, Gezhi
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 814
; CURRENT APPLICATION NUMBER: US/60/365,384
; CURRENT FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 666
; SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 181
; LENGTH: 3298
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-365-384-181

Query Match 17.7%; Score 64.5; DB 7; Length 3298;
Best Local Similarity 34.4%; Pred. No. 3.2e+02;
Matches 21; Conservative 12; Mismatches 21; Indels 7; Gaps 3;

```
Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPEKYLFQFHNLNSNLDRIFD 65
           ||  :|      :  || :| |  ||: | ||| :|:| :| |       |:|| |:: :
Db    2138 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLLN 2193

Qy      66 K 66
           :
Db    2194 E 2194
```

Search completed: October 10, 2002, 02:09:00
Job time: 1443 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:01:59 ; Search time 118.38 Seconds
(without alignments)
102.295 Million cell updates/sec

Title: US-09-289-346A-2
Perfect score: 362
Sequence: 1 TLVWGEFQVDGAAAAGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	300	82.9	351	12	Q91R10	Q91r10 tomato seve
2	298	82.3	352	12	Q9E000	Q9e000 tomato rugo
3	281	77.6	226	12	O09727	O09727 leonurus mo
4	281	77.6	226	12	Q9WHF6	Q9whf6 tomato mild
5	279	77.1	361	12	Q67574	Q67574 bean golden
6	277	76.5	225	12	Q9QDB1	Q9qdb1 cowpea gold
7	277	76.5	314	12	Q9ELT8	Q9elt8 sweet potat
8	277	76.5	364	12	Q9QS55	Q9qs55 sweet potat
9	274	75.7	185	12	Q98693	Q98693 sida golden
10	266	73.5	149	12	P88975	P88975 macroptiliu
11	266	73.5	233	12	Q9YLA4	Q9yla4 macroptiliu
12	260	71.8	234	12	O39180	O39180 geminivirid
13	257	71.0	190	12	Q9Z089	Q9z089 tobacco lea
14	257	71.0	190	12	Q9Z084	Q9z084 tobacco lea
15	256	70.7	190	12	Q9W827	Q9w827 tobacco lea
16	256	70.7	208	12	Q9Z0C4	Q9z0c4 tobacco lea
17	256	70.7	359	12	Q91M88	Q91m88 tobacco lea
18	254	70.2	208	12	Q9Z0C0	Q9z0c0 tobacco lea
19	254	70.2	208	12	Q9Z0B8	Q9z0b8 tobacco lea
20	252	69.6	208	12	Q9Z0B6	Q9z0b6 tobacco lea
21	251	69.3	203	12	Q9Z083	Q9z083 tobacco lea
22	251	69.3	363	12	O73577	O73577 cotton leaf
23	246	68.0	359	12	Q91B86	Q91b86 ageratum ye
24	246	68.0	360	12	Q9DXI0	Q9dxi0 ageratum ye
25	244	67.4	208	12	Q9Z0A0	Q9z0a0 tobacco lea
26	244	67.4	363	12	O72705	O72705 cotton leaf
27	244	67.4	363	12	O72719	O72719 cotton leaf
28	243	67.1	190	12	Q9Z086	Q9z086 tobacco lea
29	243	67.1	208	12	Q9Z0C6	Q9z0c6 tobacco lea
30	242	66.9	349	12	Q88888	Q88888 tomato pseu
31	242	66.9	362	12	O56816	O56816 chayote mos
32	241	66.6	359	12	Q9YZV4	Q9yzv4 tomato yell
33	241	66.6	359	12	Q9YZV2	Q9yzv2 tomato yell
34	241	66.6	359	12	Q9YUX7	Q9yux7 tomato yell
35	241	66.6	359	12	Q9YL27	Q9yl27 tomato yell
36	240	66.3	348	12	Q911W5	Q911w5 macroptiliu
37	240	66.3	353	12	O72692	O72692 beet curly
38	240	66.3	354	12	Q65438	Q65438 beet curly
39	240	66.3	359	12	Q88942	Q88942 tomato yell
40	239	66.0	190	12	Q9Z0A7	Q9z0a7 tobacco lea
41	239	66.0	362	12	Q9IN48	Q9in48 okra enatio
42	238	65.7	307	12	Q9IET7	Q9iet7 cotton leaf
43	238	65.7	361	12	O72723	O72723 cotton leaf
44	237	65.5	231	12	Q96620	Q96620 african tom
45	237	65.5	354	12	Q9IN42	Q9in42 south afric

ALIGNMENTS

RESULT 1
Q91R10
ID Q91R10 PRELIMINARY; PRT; 351 AA.
AC Q91R10;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE REP PROTEIN.
GN AC1.
OS Tomato severe rugose virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=158463;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MINAS GERAIS;
RA Rezende W.L., Goulart L.R., Parreira K.S., Figueiredo J.E.F.;
RT "The full-length DNA-A nucleotide sequence of a novel tomato-infecting
RT begomovirus, Tomato severe rugose virus, in Brazil.";
RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY029750; AAK50357.1; -.
SQ SEQUENCE 351 AA; 40122 MW; 87F937A4F873B6CF CRC64;

Query Match 82.9%; Score 300; DB 12; Length 351;
Best Local Similarity 78.6%; Pred. No. 3.9e-25;
Matches 55; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||||:|| :| ||||:|||||||||| ||: ||||||||:|||:||||||||||
Db     111 TIEWGEFQIDGRSARGGCQTANDAAAEALNAPSKDVALQIIREKLPEKFLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| : |||
Db     171 DRIFARAPEP 180
```

RESULT 2
Q9E000
ID Q9E000 PRELIMINARY; PRT; 352 AA.

```
AC   Q9E000;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN.
GN   AC1.
OS   Tomato rugose mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=134599;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Fernandes J.J., Fontes E.P.B., Brommonschenkel S.H., Carvalho M.G.,
RA   Zambolim E.M., Zerbini F.M.;
RT   "Molecular Cloning and Characterization of Tomato rugose mosaic virus
RT   (TRMV), a Begomovirus Isolated from Tomatoes at Triangulo Mineiro,
RT   Minas Gerais, Brazil.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF291705; AAG15546.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   352 AA;  40012 MW;  47CD55838E24D613 CRC64;

  Query Match             82.3%;  Score 298;  DB 12;  Length 352;
  Best Local Similarity   78.6%;  Pred. No. 6.5e-25;
  Matches   55;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|| :| |||||:|||||||||| ||: ||||||||:|||:||||||||||
Db     111 TIEWGEFQIDGRSARGGCQTANDAAAEALNAPSKDIALQIIREKMPEKFLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| : |||
Db     171 DRIFARAPEP 180


RESULT   3
O09727
ID   O09727      PRELIMINARY;      PRT;   226 AA.
AC   O09727;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   REP.
OS   Leonurus mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=58177;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LEMV- BRAZIL 1;
RA   Faria J.C., Maxwell D.P.;
RT   "Variability in geminivirus associated with Phaseolus vulgaris in
RT   Brazil.";
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U92532; AAB51157.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     226    226
SQ   SEQUENCE   226 AA;  25617 MW;  73CDB6E766083FC5 CRC64;

  Query Match             77.6%;  Score 281;  DB 12;  Length 226;
  Best Local Similarity   77.1%;  Pred. No. 3e-23;
  Matches   54;  Conservative    5;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||||| :: || || |||||||||| |  |||||:||:|||||||||||||||
Db     111 TVEWGEFQVDGRSSRGGQQTVNDAAAEALNAPDKRTALQIIKEKLPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180


RESULT   4
Q9WHF6
ID   Q9WHF6      PRELIMINARY;      PRT;   226 AA.
AC   Q9WHF6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   tomato mild mottle geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX   NCBI_TaxID=92943;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HN96-H5KW;
RA   Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA   Maxwell D.P.;
RT   "Molecular characterization and DNA-based detection methods for
RT   vegetable-infecting geminiviruses in Central America.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF131071; AAD33471.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     226    226
SQ   SEQUENCE   226 AA;  25941 MW;  2EA4116712871A23 CRC64;

  Query Match             77.6%;  Score 281;  DB 12;  Length 226;
  Best Local Similarity   72.9%;  Pred. No. 3e-23;
  Matches   51;  Conservative   13;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|| :| || ||:|||||||||||||||||::||:||:|:::|||:|||:|||
Db     111 TIEWGEFQIDGRSARGGQQTANDAAAEALNASSKEEAMRIIKEKLPKEFLFQYHNLSSNL 170

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180


RESULT   5
Q67574
ID   Q67574      PRELIMINARY;      PRT;   361 AA.
AC   Q67574;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE  REPLICATIVE PROTEIN.
GN   AL1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,
RA   Maxwell D.P., Russell D.R.;
RT   "Cloning of the complete DNA genomes of four bean-infecting
RT   geminiviruses and determining their infectivity by electric discharge
RT   particle acceleration.";
RL   Phytopathology 81:980-985(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,
RA   Morales F.J., Maxwell D.P.;
RT   "Differentiation of bean-infecting geminiviruses by nucleic acid
```

```
RT   hybridization probes and aspects of bean golden mosaic in Brazil.";
RL   Plant Dis. 75:336-342(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RT   "Genetic diversity in geminiviruses causing bean golden mosaic
RT   disease: The nucleotide sequence of the infectious cloned DNA
RT   components of a Brazilian isolate of bean golden mosaic geminivirus.";
RL   Submitted (MAY-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; M88686; AAA46312.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   361 AA;  41041 MW;  0094C7ACAF06B788 CRC64;


  Query Match             77.1%;  Score 279;  DB 12;  Length 361;
  Best Local Similarity   76.1%;  Pred. No. 8.4e-23;
  Matches   51;  Conservative   10;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           || ||||| :| || ||:||||:||||||||||||:|||:|||:|||:|||:||||||
Db     113 WGHFQVDGRSARGGQQTANDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
           | | |:|
Db     173 FTKAPDP 179


RESULT   6
Q9QDB1
ID   Q9QDB1      PRELIMINARY;      PRT;   225 AA.
AC   Q9QDB1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   cowpea golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=69263;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CGMV-BR;
RA   Faria J.C.;
RT   "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT   Brazil.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF188708; AAF06318.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     225    225
SQ   SEQUENCE   225 AA;  25766 MW;  1089CB6BD8D15B5D CRC64;


  Query Match             76.5%;  Score 277;  DB 12;  Length 225;
  Best Local Similarity   76.1%;  Pred. No. 8.2e-23;
  Matches   51;  Conservative    9;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           || ||:|| :| || || ||||:||||||||||||:|||:||:|||:|||:|||:||||||
Db     113 WGHFQIDGRSARGGQQTINDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
           | | |||
Db     173 FKKPPEP 179


RESULT   7
Q9ELT8
ID   Q9ELT8      PRELIMINARY;      PRT;   314 AA.
AC   Q9ELT8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION ASSOCIATION PROTEIN.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF288227; AAG01006.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   314 AA;  35153 MW;  6B6220613046943F CRC64;


  Query Match             76.5%;  Score 277;  DB 12;  Length 314;
  Best Local Similarity   79.4%;  Pred. No. 1.2e-22;
  Matches   54;  Conservative    4;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||||| :| || ||:|||||||||| ||| ||||||||:|||||||||| |||
Db     110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy      61 DRIFDKTP 68
           ||||   |
Db     170 DRIFSPPP 177


RESULT   8
Q9QS55
ID   Q9QS55      PRELIMINARY;      PRT;   364 AA.
AC   Q9QS55;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION INITIATION PROTEIN AC1.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A.;
RT   "Cloning of a DNA-A-like genomic component of sweet potato leaf curl
RT   virus : nucleotide sequence and phylogenetic relationships.";
RL   Online Publication.
DR   EMBL; AF104036; AAD47173.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   364 AA;  40680 MW;  5F79752431A09D6E CRC64;
```

```
  Query Match             76.5%;  Score 277;  DB 12;  Length 364;
  Best Local Similarity   79.4%;  Pred. No. 1.4e-22;
  Matches   54;  Conservative    4;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||||| :| || ||:|||||||||| ||| ||||||||:|||||||||| |||
Db     110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy      61 DRIFDKTP 68
           ||||   |
Db     170 DRIFSPPP 177


RESULT   9
Q98693
ID   Q98693      PRELIMINARY;      PRT;   185 AA.
AC   Q98693;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   AC1.
OS   sida golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51034;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA;
RA   Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RT   "Genetic Diversity among geminiviruses associated with the weed
RT   species Sida spp, Macroptilium lathyroides, and Wissadula amplissima
RT   from Jamaica.";
RL   Plant Dis. 81:1251-1258(1997).
DR   EMBL; U67926; AAB97865.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1       1
FT   NON_TER     185     185
SQ   SEQUENCE   185 AA;  20975 MW;  3913850A025A5EE1 CRC64;


  Query Match             75.7%;  Score 274;  DB 12;  Length 185;
  Best Local Similarity   70.0%;  Pred. No. 1.4e-22;
  Matches   49;  Conservative   13;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:|| :| || ||:||||||||||: :||:||:|||||:||:||||:|||:|||:
Db      89 TIEWGVFQIDGRSARGGQQTANDAAAEALNSGTKEDALKIIREKLPERYLFQYHNLSSNI 148

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     149 DRIFSKPPEP 158


RESULT  10
P88975
ID   P88975      PRELIMINARY;      PRT;   149 AA.
AC   P88975;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   AC1.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICAN;
RA   Roye M.E.;
RL   Thesis (1996), Biochemistry, University of the West Indies, Jamaica.
DR   EMBL; U75278; AAB36919.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1       1
FT   NON_TER     149     149
SQ   SEQUENCE   149 AA;  16785 MW;  E4CF5EED4C9CD508 CRC64;


  Query Match             73.5%;  Score 266;  DB 12;  Length 149;
  Best Local Similarity   68.6%;  Pred. No. 8.4e-22;
  Matches   48;  Conservative   13;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:|| :| || ||||||||||||: :|| |::|::||:|||:|||:|||:|||
Db      52 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 111

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     112 DRIFMKDPEP 121


RESULT  11
Q9YLA4
ID   Q9YLA4      PRELIMINARY;      PRT;   233 AA.
AC   Q9YLA4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E.;
RT   "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
RT   from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E., McLaughlin W.A., Maxwell D.P.;
RT   "Molecular characterization of two distinct geminiviruses infecting M.
RT   lathyroides from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF098940; AAD17850.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     233     233
SQ   SEQUENCE   233 AA;  26356 MW;  AA490AF4D2166A02 CRC64;


  Query Match             73.5%;  Score 266;  DB 12;  Length 233;
  Best Local Similarity   68.6%;  Pred. No. 1.4e-21;
  Matches   48;  Conservative   13;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:|| :| || ||||||||||||: :|| |::|::||:|||:|||:|||:|||
Db     110 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 169

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     170 DRIFMKDPEP 179


RESULT  12
O39180
```

```
ID   O39180      PRELIMINARY;      PRT;   234 AA.
AC   O39180;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   AC1 (FRAGMENT).
OS   Geminiviridae.
OC   Viruses; ssDNA viruses.
OX   NCBI_TaxID=10811;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOMATO STRAIN;
RA   Guzman P., Arredondo C.R., Emmatty D., Portillo R.J., Gilbertson R.L.;
RT   "Partial Characterization of Two Whitefly-Transmitted Geminiviruses
RT   Infecting Tomatoes in Venezuela.";
RL   Plant Dis. 81:312-312(1997).
DR   EMBL; AF026553; AAB82605.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     234     234
SQ   SEQUENCE   234 AA;  26486 MW;  9ED8F0697105CD19 CRC64;
```

```
  Query Match             71.8%;  Score 260;  DB 12;  Length 234;
  Best Local Similarity   67.1%;  Pred. No. 6.5e-21;
  Matches   47;  Conservative   14;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:|| :| || ||:||||||||||: :||||::||:||:|||:|||:|||:|:|
Db     110 TIEWGLFQIDGRSARGGQQTANDAAAEALNSGTKEEAMKIIKEKLPEKFLFQYHNLSSDL 169

Qy      61 DRIFDKTPEP 70
           |||| | | |
Db     170 DRIFMKAPNP 179
```

```
RESULT  13
Q9Z089
ID   Q9Z089      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z089;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA3-1, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001315; BAA34033.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1       1
FT   NON_TER     190     190
SQ   SEQUENCE   190 AA;  21432 MW;  AAC093D1D1610FAD CRC64;
```

```
  Query Match             71.0%;  Score 257;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 1.1e-20;
  Matches   50;  Conservative    9;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || || ||:|| :| |||| :||| |||||||||| ||| |||||:|: ::||:|||||||
```

```
Db      85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFTASSFDQVPE 169
```

```
RESULT  14
Q9Z084
ID   Q9Z084      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z084;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA5-2, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA5;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001318; BAA34039.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1       1
FT   NON_TER     190     190
SQ   SEQUENCE   190 AA;  21444 MW;  AAC1C2943E3F01AD CRC64;
```

```
  Query Match             71.0%;  Score 257;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 1.1e-20;
  Matches   50;  Conservative    9;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || || ||:|| :| |||| :||| |||||||||| ||| |||||:|: ::||:|||||||
Db      85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFTASSFDQVPE 169
```

```
RESULT  15
Q9W827
ID   Q9W827      PRELIMINARY;      PRT;   190 AA.
AC   Q9W827;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE GORA1-1, PARTIAL AND
DE   COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GORA1;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001303; BAA34010.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
```

```
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  93C3742A8EBDB7EB CRC64;


  Query Match             70.7%;  Score 256;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 1.4e-20;
  Matches   50;  Conservative    9;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || || ||||| :| |||| :||| |||||||:|| ||| |||||:|: ::||:|||||||
Db      85 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFSNSSFDQVPE 169


Search completed: October 10, 2002, 02:01:59
Job time: 1092 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:48:42 ; Search time 69.63 Seconds
(without alignments)
96.600 Million cell updates/sec

Title: US-09-289-346A-2
Perfect score: 362
Sequence: 1 TLVWGEFQVDGAAAAGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	349	96.4	352	1	QQCVL1	AL1 protein - toma
2	249	68.8	361	1	QQCVPT	AL1 protein - pota
3	236	65.2	358	2	S07594	hypothetical prote
4	229	63.3	362	1	JQ1887	AL1 protein - toma
5	222	61.3	359	2	S39211	gene C1 protein -
6	218	60.2	349	2	JQ2300	replicase - pepper
7	218	60.2	349	2	S31875	AL1 protein - pepp
8	215	59.4	359	2	S22593	hypothetical prote
9	212	58.6	351	2	JQ2327	AL1 protein - Indi
10	212	58.6	355	1	QQCVW1	AV1 protein - abut
11	209	57.7	358	1	JQ1870	AL1 protein - toma
12	209	57.7	359	2	S39235	gene C1 protein -
13	209	57.7	385	2	S28360	AL1 protein - beet
14	202	55.8	360	2	S59885	replication-associ
15	200	55.2	357	1	QQCVC1	AL1 protein - toma
16	142	39.2	131	2	S45059	AC1 protein (clone
17	124	34.3	347	1	QQCVS1	AL1 protein - squa
18	68.5	18.9	1713	2	A55347	adhesive ligand ep
19	65.5	18.1	1610	2	A46227	voltage-dependent
20	65.5	18.1	1646	2	JH0422	voltage-dependent
21	65.5	18.1	2161	2	JH0564	calcium channel al
22	65.5	18.1	2181	2	A38198	calcium channel al
23	65.5	18.1	2203	2	T42742	voltage-dependent
24	65	18.0	371	2	B69502	alcohol dehydrogen
25	64.5	17.8	299	2	B71967	probable peptidyl-
26	64.5	17.8	335	1	DEBSGF	glyceraldehyde-3-p
27	64.5	17.8	557	2	A47162	thiolesterase B (E
28	63.5	17.5	392	2	T45290	acetyl-CoA C-acety
29	63	17.4	397	2	B71078	probable NADH oxid
30	62.5	17.3	419	2	S73846	MG223 homolog F10_
31	62.5	17.3	1033	2	E97700	hypothetical prote
32	62	17.1	154	2	AD3475	nitrogen regulator
33	62	17.1	181	2	G97976	conserved hypothet
34	62	17.1	295	2	D42452	C1 protein - tobac
35	62	17.1	308	2	C84072	UDP-glucose 4-epim
36	62	17.1	1502	1	RGBYH1	CYC1/CYP3 transcri
37	61.5	17.0	481	2	A70091	probable phosphoes
38	61	16.9	338	2	AG3582	iron(III)-binding
39	61	16.9	392	2	T44362	acetyl-CoA C-acety
40	61	16.9	587	2	JC1419	Fc gamma (IgG) rec
41	60.5	16.7	136	2	T22240	hypothetical prote
42	60.5	16.7	201	2	A81380	probable flagellar
43	60.5	16.7	652	2	E89841	fructose specific
44	60	16.6	384	1	W2WL41	E2 protein - human
45	60	16.6	447	2	S52437	CDP-diacylglycerol

ALIGNMENTS

RESULT 1
QQCVL1
AL1 protein - tomato golden mosaic virus
C;Species: tomato golden mosaic virus
A;Note: host Nicotiana sp. (tobacco)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C;Accession: A04170
R;Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.
EMBO J. 3, 2197-2205, 1984
A;Title: Complete nucleotide sequence of the infectious cloned DNA components of toma
A;Reference number: A04163
A;Accession: A04170
A;Molecule type: DNA
A;Residues: 1-352 <HAM>
C;Comment: The genome consists of two circular, single-stranded DNA components, DNA A
.
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 96.4%; Score 349; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 1.5e-30;
Matches 67; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||| :| ||||||||||||||||||||||||||||||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
QQCVPT
AL1 protein - potato yellow mosaic virus (isolate Venezuela)
C;Species: potato yellow mosaic virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C;Accession: JU0364
R;Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.
J. Gen. Virol. 72, 1515-1520, 1991
A;Title: The nucleotide sequence of the infectious cloned DNA components of potato ye
A;Reference number: JU0362; MUID:91311403
A;Accession: JU0364
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-361 <COU>
A;Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             68.8%;  Score 249;  DB 1;  Length 361;
  Best Local Similarity   66.7%;  Pred. No. 1.4e-19;
  Matches   46;  Conservative   12;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:|| :| || || |||||||||: :|| |::||:||:|||:|||:|||: ||
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 DRIFDKTPE 69
           |||| | ||
Db     170 DRIFMKAPE 178


RESULT   3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C;Species: cassava latent virus
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C;Accession: S07594
R;Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A;Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A;Reference number: S07590; MUID:90174930
A;Accession: S07594
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <MOR>
A;Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C;Genetics:
A;Map position: segment DNA1
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             65.2%;  Score 236;  DB 2;  Length 358;
  Best Local Similarity   60.0%;  Pred. No. 3.6e-18;
  Matches   42;  Conservative   14;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|| :| || |::||| |:||: || ||| :||| :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   4
JQ1887
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1887
R;Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A;Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A;Reference number: JQ1885; MUID:93139778
A;Accession: JQ1887
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-362 <DRY>
A;Cross-references: GB:S53251
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             63.3%;  Score 229;  DB 1;  Length 362;
  Best Local Similarity   52.9%;  Pred. No. 2.1e-17;
  Matches   45;  Conservative   10;  Mismatches   14;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || |||||:|| :| || |::||| |:|||  || ||| ::||  |: |: |||||||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                  ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   5
S39211
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C;Accession: S39211
R;Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A;Description: High similarity among the tomato yellow leaf curl virus isolates from
A;Reference number: S39209
A;Accession: S39211
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <NOR>
A;Cross-references: EMBL:Z25751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             61.3%;  Score 222;  DB 2;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 1.2e-16;
  Matches   40;  Conservative   10;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|| :| || ||:||| |:|:|| || ||| :|:|  |  |:  |||:||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   6
JQ2300
replicase - pepper huasteco virus (component A)
N;Alternate names: ORF AL1 protein
C;Species: pepper huasteco virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C;Accession: JQ2300
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A;Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A;Reference number: JQ2299; MUID:94015007
A;Accession: JQ2300
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.2%;  Score 218;  DB 2;  Length 349;
  Best Local Similarity   57.1%;  Pred. No. 3.3e-16;
  Matches   40;  Conservative   14;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|| :| || |::||  |:|||::| |||||||:|: |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   7
S31875
AL1 protein - pepper rizado amarillo virus
C;Species: pepper rizado amarillo virus
```

```
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Sep-1999
C;Accession: S31875
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera-Bu
submitted to the EMBL Data Library, February 1993
A;Description: Complete nucleotide sequence of pepper huasteco virus: analysis and compa
A;Reference number: S31872
A;Accession: S31875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
A;Note: the source is designated as pepper huasteco virus
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.2%;  Score 218;  DB 2;  Length 349;
  Best Local Similarity   57.1%;  Pred. No. 3.3e-16;
  Matches   40;  Conservative   14;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|| :| || |::||  |:|||::| |||||||:|: |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   8
S22593
hypothetical protein C4 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Sep-1999
C;Accession: S22593
R;Kheyr-Pour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.
Nucleic Acids Res. 19, 6763-6769, 1991
A;Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monoparti
A;Reference number: S22588; MUID:92107660
A;Accession: S22593
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-359 <KHE>
A;Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.4%;  Score 215;  DB 2;  Length 359;
  Best Local Similarity   55.1%;  Pred. No. 7.2e-16;
  Matches   38;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|| :| || ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   9
JQ2327
AL1 protein - Indian cassava mosaic virus
N;Alternate names: replication-associated protein
C;Species: Indian cassava mosaic virus
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C;Accession: JQ2327; S35883
R;Hong, Y.G.; Robinson, D.J.; Harrison, B.D.
J. Gen. Virol. 74, 2437-2443, 1993
A;Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-tran
A;Reference number: JQ2326; MUID:94065670
A;Accession: JQ2327
A;Molecule type: DNA
A;Residues: 1-351 <HON>
A;Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             58.6%;  Score 212;  DB 2;  Length 351;
  Best Local Similarity   59.7%;  Pred. No. 1.5e-15;
  Matches   40;  Conservative   10;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           || ||:|| :| || |::||| | |||: || |||:|:|| |  ||  ||:::||||||
Db     113 WGTFQIDGRSARGGQQSANDAYAAALNSGSKSEALKILRELAPRDYLRDFHHISSNLDRI 172

Qy      64 FDKTPEP 70
           | | | |
Db     173 FTKPPPP 179


RESULT  10
QQCVW1
AV1 protein - abutilon mosaic virus (isolate West India)
C;Species: abutilon mosaic virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Apr-1994
C;Accession: A36214
R;Frischmuth, T.; Zimmat, G.; Jeske, H.
Virology 178, 461-468, 1990
A;Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as
A;Reference number: A36214; MUID:91020984
A;Accession: A36214
A;Molecule type: DNA
A;Residues: 1-355 <FRI>
A;Cross-references: EMBL:X15983
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             58.6%;  Score 212;  DB 1;  Length 355;
  Best Local Similarity   55.7%;  Pred. No. 1.5e-15;
  Matches   39;  Conservative   13;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |  |||||:|| :| || ||:||: |:||||    : || |::|: |: |: | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
JQ1870
AL1 protein - tomato mottle virus (isolate Florida)
C;Species: tomato mottle virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1870
R;Abouzid, A.M.; Polston, J.E.; Hiebert, E.
J. Gen. Virol. 73, 3225-3229, 1992
A;Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated f
A;Reference number: JQ1869; MUID:93107858
A;Accession: JQ1870
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <ABO>
A;Cross-references: GB:L14460
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             57.7%;  Score 209;  DB 1;  Length 358;
  Best Local Similarity   52.9%;  Pred. No. 3.3e-15;
  Matches   37;  Conservative   17;  Mismatches   16;  Indels    0;  Gaps    0;
```

```
Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|| :| || |::||: |:|||| | : || ::||: |: :: | ||: |||
Db     107 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 166

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     167 ERIFAKAPEP 176
```

RESULT 12
S39235
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Sep-1997
C;Accession: S39235
R;Crespi, S.; Noris, E.; Vaira, A.; Bosco, D.; Accotto, G.
submitted to the EMBL Data Library, December 1993
A;Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.
A;Reference number: S39233
A;Accession: S39235
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <CRE>
A;Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             57.7%;  Score 209;  DB 2;  Length 359;
  Best Local Similarity   53.6%;  Pred. No. 3.3e-15;
  Matches   37;  Conservative   13;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|| :| || ||:||| |:|:|| || ||| :|:: |  |:  |||::||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINARSKSEALDVIKQLAPRDYVLHFHNISSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179
```

RESULT 13
S28360
AL1 protein - beet curly top virus
C;Species: beet curly top virus
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
C;Accession: S28360
R;Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
EMBO J. 5, 1761-1767, 1986
A;Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly to
A;Reference number: S28360
A;Accession: S28360
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-385 <STA>
A;Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             57.7%;  Score 209;  DB 2;  Length 385;
  Best Local Similarity   52.9%;  Pred. No. 3.5e-15;
  Matches   37;  Conservative   17;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|| :| || ||:||: |:||||:| ::|||||::|: |: |  | ||| :|
Db     137 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 196

Qy      61 DRIFDKTPEP 70
           :|| : |:|
Db     197 QKIFQRPPDP 206
```

RESULT 14
S59885
replication-associated protein C1 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
C;Accession: S59885
R;Hong, Y.; Harrison, B.D.
submitted to the EMBL Data Library, February 1995
A;Description: Nucleotide sequences from tomato leaf curl viruses from different coun
d geminiviruses.
A;Reference number: S58346
A;Accession: S59885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <HON>
A;Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA88229.1; PID:g974211
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             55.8%;  Score 202;  DB 2;  Length 360;
  Best Local Similarity   57.6%;  Pred. No. 1.9e-14;
  Matches   38;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           :| ||:|| :| || |::||| |||:|: ||  || |:||| |: :: ||||||:|||||
Db     113 FGVFQIDGRSARGGQQSANDAYAEAINSGSKAAALDILREKAPKDFVLQFHNLNANLDRI 172

Qy      64 FDKTPE 69
           |  : |
Db     173 FTPSAE 178
```

RESULT 15
QQCVC1
AL1 protein - tomato yellow leaf curl virus
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C;Accession: D40779
R;Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.
Virology 185, 151-161, 1991
A;Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin
A;Reference number: A40779; MUID:92024070
A;Accession: D40779
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-357 <NAV>
A;Cross-references: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             55.2%;  Score 200;  DB 1;  Length 357;
  Best Local Similarity   62.3%;  Pred. No. 3.1e-14;
  Matches   38;  Conservative   11;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           :|  |:|| :| || |::||| |||||: || ||| |::|| |: |: ||||||:||||||
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171
```

Search completed: October 10, 2002, 01:48:42
Job time: 560 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:47:17 ; Search time 133.15 Seconds
(without alignments)
58.394 Million cell updates/sec

Title: US-09-289-346A-2
Perfect score: 362
Sequence: 1 TLVWGEFQVDGAAAAGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	362	100.0	70	21	AAB18678	Mutant peptide der
2	349	96.4	70	21	AAB18677	Peptide fragment f
3	349	96.4	356	21	AAB18687	Amino acid sequenc
4	341	94.2	70	21	AAB18685	Mutant peptide der
5	338	93.4	70	21	AAB18688	Mutant peptide der
6	337	93.1	70	21	AAB18692	Mutant peptide der
7	335	92.5	70	21	AAB18684	Mutant peptide der
8	335	92.5	70	21	AAB18690	Mutant peptide der
9	334	92.3	70	21	AAB18686	Mutant peptide der
10	333	92.0	70	21	AAB18689	Mutant peptide der
11	331	91.4	70	21	AAB18680	Mutant peptide der
12	331	91.4	70	21	AAB18691	Mutant peptide der
13	328	90.6	70	21	AAB18681	Mutant peptide der
14	327	90.3	70	21	AAB18683	Mutant peptide der
15	325	89.8	70	21	AAB18682	Mutant peptide der
16	317	87.6	70	21	AAB18679	Mutant peptide der
17	215	59.4	353	18	AAW34338	Bean golden mosaic
18	215	59.4	353	18	AAW34332	Bean golden mosaic
19	215	59.4	353	18	AAW34333	Bean golden mosaic
20	215	59.4	353	18	AAW34334	Bean golden mosaic
21	215	59.4	353	18	AAW34335	Sardinian tomato y
22	215	59.4	359	17	AAR88870	Sardinian tomato y
23	215	59.4	359	17	AAR88871	Sardinian tomato y
24	215	59.4	359	17	AAR88872	ORF 4 gene product
25	213	58.8	353	8	AAP70407	Tomato mottle viru
26	213	58.8	361	18	AAW34336	Tomato mottle viru
27	213	58.8	361	18	AAW34324	Tomato mottle viru
28	213	58.8	361	18	AAW34325	Tomato mottle viru
29	213	58.8	361	18	AAW34326	Product of ORF 4 f
30	206.5	57.0	361	8	AAP70562	Tobacco leaf curli
31	201	55.5	362	19	AAW56495	Tomato yellow leaf
32	200	55.2	357	18	AAW34329	Tomato yellow leaf
33	200	55.2	357	18	AAW34330	Tomato yellow leaf
34	200	55.2	357	18	AAW34331	Tomato yellow leaf
35	192	53.0	357	18	AAW34337	Drosophila melanog
36	81.5	22.5	945	22	ABB65231	HIV-1 strain YBF30
37	69.5	19.2	512	19	AAW68473	Human laminin 5 po
38	68.5	18.9	1693	21	AAB48457	Human laminin 5 po
39	68.5	18.9	1693	21	AAB48459	Deduced sequence o
40	68.5	18.9	1713	16	AAR70148	Human laminin 5 po
41	68.5	18.9	1713	21	AAB48458	Human laminin 5 po
42	68.5	18.9	1724	21	AAB48456	Drosophila melanog
43	66.5	18.4	665	22	ABB70991	Human calcium chan
44	65.5	18.1	619	13	AAR27651	Sequence of the al
45	65.5	18.1	2161	14	AAR33545	

ALIGNMENTS

RESULT 1
AAB18678
ID AAB18678 standard; peptide; 70 AA.
XX
AC AAB18678;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 12
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 13
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 15
FT /note= "wild type residue replaced with Ala"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.

XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DR WPI; 2000-618851/59.
XX
PT Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT -
XX
PS Claim 53; Page 42-43; 73pp; English.
XX
CC The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
SQ Sequence 70 AA;

Query Match 100.0%; Score 362; DB 21; Length 70;
Best Local Similarity 100.0%; Pred. No. 4e-37;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 tlvwgefqvdgaaaaggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy 61 DRIFDKTPEP 70
||||||||||
Db 61 drifdktpep 70

RESULT 2
AAB18677
ID AAB18677 standard; peptide; 70 AA.
XX
AC AAB18677;
XX
DT 22-JAN-2001 (first entry)
XX
DE Peptide fragment from Rep protein of TGMV (amino acids 110-179).
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Tomato golden mosaic virus.
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DR WPI; 2000-618851/59.
XX
PT Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT -
XX
PS Disclosure; Page 18; 73pp; English.
XX
CC The present sequence is derived from a geminivirus replication (Rep)
CC protein, which is also known as AL1. AL1 binds double-stranded DNA,
CC catalyses cleavage and ligation of single-stranded DNA, and interacts
CC with other viral and host proteins. Mutants of the AL1 protein are used
CC to produce transgenic plants. The mutation in AL1 is present in a
CC ribosome binding region, and expression of mutant AL1 protein imparts
CC increased resistance to geminivirus infection in the plant. Mutant AL1
CC proteins are useful for producing plants having increased resistance or
CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,
CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl
CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato
CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,
CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or
CC beet curly top virus.
XX
SQ Sequence 70 AA;

Query Match 96.4%; Score 349; DB 21; Length 70;
Best Local Similarity 95.7%; Pred. No. 1.6e-35;
Matches 67; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
|||||||||||| :| ||||||||||||||||||||||||||||||||||||||||||||||
Db 1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy 61 DRIFDKTPEP 70
||||||||||
Db 61 drifdktpep 70

RESULT 3
AAB18687
ID AAB18687 standard; peptide; 356 AA.
XX
AC AAB18687;
XX
DT 22-JAN-2001 (first entry)
XX
DE Amino acid sequence of a geminivirus replication protein of TGMV.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 354
FT /note= "unspecified amino acid"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DR WPI; 2000-618851/59.
XX
PT Transgenic plants with increased resistance to geminivirus infection

```
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 47-48; 73pp; English.
XX
CC   The present sequence represents a geminivirus replication (Rep)
CC   protein, which is also known as AL1. AL1 binds double-stranded DNA,
CC   catalyses cleavage and ligation of single-stranded DNA, and interacts
CC   with other viral and host proteins. Mutants of the AL1 protein are used
CC   to produce transgenic plants. The mutation in AL1 is present in a
CC   ribosome binding region, and expression of mutant AL1 protein imparts
CC   increased resistance to geminivirus infection in the plant. Mutant AL1
CC   proteins are useful for producing plants having increased resistance or
CC   reduced sensitivity to a geminivirus such as tomato golden mosaic virus,
CC   tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl
CC   virus, African cassava mosaic virus, Indian cassava mosaic virus, potato
CC   yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,
CC   squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or
CC   beet curly top virus.
XX
SQ   Sequence   356 AA;


  Query Match             96.4%;  Score 349;  DB 21;  Length 356;
  Best Local Similarity   95.7%;  Pred. No. 1.2e-34;
  Matches   67;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||| :| ||||||||||||||||||||||||||||||||||||||||||||
Db     110 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 169

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     170 drifdktpep 179


RESULT   4
AAB18685
ID   AAB18685 standard; peptide; 70 AA.
XX
AC   AAB18685;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 10
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 46; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             94.2%;  Score 341;  DB 21;  Length 70;
  Best Local Similarity   94.3%;  Pred. No. 1.6e-34;
  Matches   66;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||| | :| ||||||||||||||||||||||||||||||||||||||||||||
Db       1 tlvwgefqvagrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT   5
AAB18688
ID   AAB18688 standard; peptide; 70 AA.
XX
AC   AAB18688;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 19
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 20
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
```

```
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 48; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             93.4%;  Score 338;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 3.6e-34;
  Matches   65;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |||||||||||| :| |||  ||||||||||||||||||||||||||||||||||||||||
Db      1 tlvwgefqvdgrsarggcaasndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   6
AAB18692
ID   AAB18692 standard; peptide; 70 AA.
XX
AC   AAB18692;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 66
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 69
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
```

```
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 50; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             93.1%;  Score 337;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 4.8e-34;
  Matches   65;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |||||||||||| :| ||||||||||||||||||||||||||||||||||||||||||||
Db      1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||| || |
Db     61 drifdatpap 70


RESULT   7
AAB18684
ID   AAB18684 standard; peptide; 70 AA.
XX
AC   AAB18684;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 7
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 8
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
```

```
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             92.5%;  Score 335;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 8.5e-34;
  Matches   65;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||  ||| :| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1 tlvwgeaavdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT   8
AAB18690
ID   AAB18690 standard; peptide; 70 AA.
XX
AC   AAB18690;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 27
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 30
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             92.5%;  Score 335;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 8.5e-34;
  Matches   65;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||| :| |||||||||||| || ||||||||||||||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaaalaasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT   9
AAB18686
ID   AAB18686 standard; peptide; 70 AA.
XX
AC   AAB18686;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 24
FT                   /note= "wild type residue replaced with Leu"
FT   Misc-difference 25
FT                   /note= "wild type residue replaced with Leu"
FT   Misc-difference 26
FT                   /note= "wild type residue replaced with Leu"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
```

```
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 46; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;
```

```
  Query Match             92.3%;  Score 334;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 1.1e-33;
  Matches   64;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||| :| ||||||||   ||||||||||||||||||||||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndlllealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70
```

```
RESULT  10
AAB18689
ID   AAB18689 standard; peptide; 70 AA.
XX
AC   AAB18689;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 22
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 23
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
```

```
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 48-49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;
```

```
  Query Match             92.0%;  Score 333;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 1.5e-33;
  Matches   65;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||| :| ||||||  ||||||||||||||||||||||||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsaaaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70
```

```
RESULT  11
AAB18680
ID   AAB18680 standard; peptide; 70 AA.
XX
AC   AAB18680;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 42
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 43
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 44
```

```
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 43-44; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             91.4%;  Score 331;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 2.7e-33;
  Matches   64;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||  :| ||||||||||||||||||||||||   ||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiiaaaipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  12
AAB18691
ID   AAB18691 standard; peptide; 70 AA.
XX
AC   AAB18691;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 34
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 35
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 36
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             91.4%;  Score 331;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 2.7e-33;
  Matches   64;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||  :| ||||||||||||||||||   ||||||||||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnassaaaalqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  13
AAB18681
ID   AAB18681 standard; peptide; 70 AA.
XX
AC   AAB18681;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
```

```
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 47
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 48
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 49
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 44; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             90.6%;  Score 328;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 6.2e-33;
  Matches   64;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||| :| |||||||||||||||||||||||||||||||   ||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipaaalfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  14
AAB18683
ID   AAB18683 standard; peptide; 70 AA.
XX
AC   AAB18683;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 59
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 61
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 62
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             90.3%;  Score 327;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 8.3e-33;
  Matches   64;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||| :| |||||||||||||||||||||||||||||||||||||||||| |
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsal 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 aaifdktpep 70


RESULT  15
AAB18682
ID   AAB18682 standard; peptide; 70 AA.
XX
AC   AAB18682;
XX
```

```
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 52
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 54
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 55
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 44-45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             89.8%;  Score 325;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 1.5e-32;
  Matches   64;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||| :| |||||||||||||||||||||||||||||||||| |  |||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfafaalnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


Search completed: October 10, 2002, 01:47:18
```

Job time: 527 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:09:50 ; Search time 34.62 Seconds
(without alignments)
78.289 Million cell updates/sec

Title: US-09-289-346A-3
Perfect score: 358
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	332	92.7	352	1	VAL1_TGMV	P03567 tomato gold
2	243	67.9	361	1	VAL1_PYMVV	P27258 potato yell
3	230	64.2	358	1	VAL1_CLVK	P14982 cassava lat
4	230	64.2	358	1	VAL1_CLVN	P14972 cassava lat
5	222	62.0	362	1	VAL1_TYLCA	P36279 tomato yell
6	217	60.6	355	1	VAL1_ABMVW	P21947 abutilon mo
7	216	60.3	349	1	VAL1_PHUV	Q06923 pepper huas
8	216	60.3	359	1	VAL1_TYLCU	P38609 tomato yell
9	215	60.1	353	1	VAL1_BGMV	P05175 bean golden
10	214	59.8	361	1	VAL1_TMOV	Q06657 tomato mott
11	212	59.2	359	1	VAL1_TYLCM	P27260 tomato yell
12	208	58.1	358	1	VAL1_BCTV	P14991 beet curly
13	190	53.1	357	1	VAL1_TYLCV	P27259 tomato yell
14	115	32.1	347	1	VAL1_SLCV	P29048 squash leaf
15	69	19.3	577	1	T2FA_DROME	Q05913 drosophila
16	66	18.4	207	1	IL6_MARMO	O35736 marmota mon
17	66	18.4	508	1	CROC_DROME	P32027 drosophila
18	65	18.2	671	1	HMOC_DROME	P22810 drosophila
19	62.5	17.5	1077	1	HLES_DROME	Q02308 drosophila
20	60.5	16.9	829	1	E74A_DROME	P20105 drosophila
21	60.5	16.9	883	1	E74B_DROME	P11536 drosophila
22	60	16.8	314	1	PMXB_HUMAN	Q99453 homo sapien
23	60	16.8	314	1	PMXB_MOUSE	O35690 mus musculu
24	59.5	16.6	470	1	RHBA_RHIME	Q9z3r2 rhizobium m
25	59	16.5	379	1	CYB_ASTPE	Q33818 asterina pe
26	59	16.5	498	1	CALR_PIG	P25117 sus scrofa
27	59	16.5	513	1	HEM0_CHICK	P18080 gallus gall
28	59	16.5	3591	1	FHAB_BORPE	P12255 bordetella
29	58.5	16.3	1264	1	MOLR_ECOLI	P33345 escherichia
30	58.5	16.3	1713	1	LMA3_HUMAN	Q16787 homo sapien
31	58	16.2	316	1	MBL_DROME	O16011 drosophila
32	58	16.2	732	1	CYG4_HUMAN	P33402 homo sapien
33	58	16.2	1081	1	UL52_HSVEB	P28962 equine herp
34	58	16.2	1992	1	TRIB_HUMAN	Q14669 homo sapien
35	57.5	16.1	303	1	ALB3_MAIZE	P10593 zea mays (m
36	57.5	16.1	338	1	HXDB_HUMAN	P31277 homo sapien
37	57.5	16.1	557	1	SASB_ANAPL	Q04791 anas platyr
38	57.5	16.1	593	1	DRI1_HUMAN	Q99856 homo sapien
39	57.5	16.1	601	1	DRI1_MOUSE	Q62431 mus musculu
40	57.5	16.1	1556	1	PRO_DROVI	Q9u6a1 drosophila
41	57	15.9	387	1	VE2_HPV41	P27552 human papil
42	57	15.9	401	1	HME1_MOUSE	P09065 mus musculu
43	57	15.9	3080	1	POLG_ZYMVC	P18479 z genome po
44	56.5	15.8	579	1	KCN2_HUMAN	Q9h2s1 homo sapien
45	56.5	15.8	664	1	PLB1_YEAST	P39105 saccharomyc

ALIGNMENTS

```
RESULT   1
VAL1_TGMV
ID   VAL1_TGMV      STANDARD;      PRT;   352 AA.
AC   P03567;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AC1.
OS   Tomato golden mosaic virus (TGMV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10831;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;
RT   "Complete nucleotide sequence of the infectious cloned DNA components
RT   of tomato golden mosaic virus: potential coding regions and regulatory
RT   sequences.";
RL   EMBO J. 3:2197-2205(1984).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K02029; -; NOT_ANNOTATED_CDS.
DR   PIR; A04170; QQCVL1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     223    230       ATP (POTENTIAL).
SQ   SEQUENCE   352 AA;  40332 MW;  C33C938E9644B4A4 CRC64;

  Query Match          92.7%;  Score 332;  DB 1;  Length 352;
  Best Local Similarity  92.9%;  Pred. No. 3.9e-31;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||||||||||||||||| || |||   ||||||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180


RESULT   2
VAL1_PYMVV
ID   VAL1_PYMVV     STANDARD;      PRT;   361 AA.
```

```
AC   P27258;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 protein.
OS   Potato yellow mosaic virus (isolate Venezuela).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10828;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91311403; PubMed=1856690;
RA   Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT   "The nucleotide sequence of the infectious cloned DNA components of
RT   potato yellow mosaic virus.";
RL   J. Gen. Virol. 72:1515-1520(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; D00940; BAA00782.1; -.
DR   PIR; JU0364; QQCVPT.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   361 AA;  40850 MW;  5627A33BF1264383 CRC64;
```

```
  Query Match             67.9%;  Score 243;  DB 1;  Length 361;
  Best Local Similarity   66.7%;  Pred. No. 8e-21;
  Matches   46;  Conservative   11;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:|||||||| || ||||| || : :  ||::||:||:|||:|||:|||: ||
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 DRIFDKTPE 69
             |||| | ||
Db     170 DRIFMKAPE 178
```

```
RESULT   3
VAL1_CLVK
ID   VAL1_CLVK      STANDARD;      PRT;   358 AA.
AC   P14982;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (40.4 kDa protein).
GN   AC1.
OS   Cassava latent virus (strain West Kenyan 844).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10818;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Gay M.R.;
RT   "Nucleotide sequence of cassava latent virus DNA.";
RL   Nature 301:260-262(1983).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; J02057; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40346 MW;  ED173E753EE92D69 CRC64;
```

```
  Query Match             64.2%;  Score 230;  DB 1;  Length 358;
  Best Local Similarity   60.0%;  Pred. No. 2.5e-19;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||| | || : | : || :||| :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 DRIFDKTPEP 70
             |||| : | |
Db     169 DRIFQEPPAP 178
```

```
RESULT   4
VAL1_CLVN
ID   VAL1_CLVN      STANDARD;      PRT;   358 AA.
AC   P14972;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (40.4 kDa protein).
GN   AC1.
OS   Cassava latent virus (strain Nigerian).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10819;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90174930; PubMed=2308831;
RA   Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT   "Nucleotide sequence of the infectious cloned DNA components of
RT   African cassava mosaic virus (Nigerian strain).";
RL   Nucleic Acids Res. 18:197-198(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X17095; CAA34953.1; -.
DR   PIR; S07594; S07594.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40435 MW;  1DB16BB0CB2D5E2C CRC64;
```

```
  Query Match             64.2%;  Score 230;  DB 1;  Length 358;
  Best Local Similarity   60.0%;  Pred. No. 2.5e-19;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||| | || : | : || :||| :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168
```

```
Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   5
VAL1_TYLCA
ID   VAL1_TYLCA     STANDARD;      PRT;   362 AA.
AC   P36279;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=36447;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93139778; PubMed=8423446;
RA   Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;
RT   "Nucleotide sequence and genome organization of tomato leaf curl
RT   geminivirus.";
RL   J. Gen. Virol. 74:147-151(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
DR   PIR; JQ1887; JQ1887.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   362 AA;  41197 MW;  343E7184B4704098 CRC64;


  Query Match             62.0%;  Score 222;  DB 1;  Length 362;
  Best Local Similarity   52.9%;  Pred. No. 2.2e-18;
  Matches   45;  Conservative    9;  Mismatches   15;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
             || ||||||:|||||||| |::||| | ||   | : || ::|| |: |: |||||||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI---------------FDKTPE 69
           |||               ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   6
VAL1_ABMVW
ID   VAL1_ABMVW     STANDARD;      PRT;   355 AA.
AC   P21947;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AC1.
OS   Abutilon mosaic virus (isolate West India).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10816;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91020984; PubMed=2219703;
RA   Frischmuth T., Zimmat G., Jeske H.;
RT   "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT   as well as eukaryotic features.";
RL   Virology 178:461-468(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR   PIR; A36214; QQCVW1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   355 AA;  40257 MW;  16A2CA8A63251E95 CRC64;


  Query Match             60.6%;  Score 217;  DB 1;  Length 355;
  Best Local Similarity   58.6%;  Pred. No. 8.1e-18;
  Matches   41;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |  |||||:|||||||| ||:||: | || |    :|| |::|: |: |: | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT   7
VAL1_PHUV
ID   VAL1_PHUV      STANDARD;      PRT;   349 AA.
AC   Q06923;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AL1.
OS   Pepper huasteco virus (PHV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=28349;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94015007; PubMed=8409944;
RA   Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA   Rivera-Bustamante R.F.;
RT   "Complete nucleotide sequence of pepper huasteco virus: analysis and
RT   comparison with bipartite geminiviruses.";
RL   J. Gen. Virol. 74:2225-2231(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70418; CAA49856.1; -.
DR   PIR; S31875; S31875.
DR   PIR; JQ2300; JQ2300.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (BY SIMILARITY).
SQ   SEQUENCE   349 AA;  39722 MW;  D5F4E76CD56370F4 CRC64;


  Query Match             60.3%;  Score 216;  DB 1;  Length 349;
  Best Local Similarity   58.6%;  Pred. No. 1e-17;
```

```
  Matches   41;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| |::||  | || ::||  |||||:|: |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   8
VAL1_TYLCU
ID   VAL1_TYLCU      STANDARD;      PRT;   359 AA.
AC   P38609;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=37139;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94256836; PubMed=8198442;
RA   Noris E., Hidalgo E., Accotto G., Moriones E.;
RT   "High similarity among the tomato yellow leaf curl virus isolates
RT   from the west Mediterranean basin: the nucleotide sequence of an
RT   infectious clone from Spain.";
RL   Arch. Virol. 135:165-170(1994).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z25751; CAA81026.1; -.
DR   PIR; S39211; S39211.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  41065 MW;  2D170A51EF80A3EC CRC64;


  Query Match             60.3%;  Score 216;  DB 1;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 1.1e-17;
  Matches   40;  Conservative    9;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| | |: | | : || :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   9
VAL1_BGMV
ID   VAL1_BGMV      STANDARD;      PRT;   353 AA.
AC   P05175;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   AL1 protein (40.2 kDa protein).
GN   AC1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Howarth A.J., Caton J., Bossert M., Goodman R.M.;
RT   "Nucleotide sequence of bean golden mosaic virus and a model for gene
RT   regulation in geminiviruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M10070; AAA46318.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   353 AA;  40190 MW;  80FA779DF6029A34 CRC64;


  Query Match             60.1%;  Score 215;  DB 1;  Length 353;
  Best Local Similarity   58.6%;  Pred. No. 1.4e-17;
  Matches   41;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||||||||||| |::||: | || | |  :|| |::|: |: |: | ||: |||
Db     110 TIEWGQFQVDGRSARGGQQSANDSYAKALNADSIESALTILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFFKVPEP 179


RESULT  10
VAL1_TMOV
ID   VAL1_TMOV      STANDARD;      PRT;   361 AA.
AC   Q06657;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AL1.
OS   Tomato mottle virus (isolate Florida) (TMoV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=36449;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93107858; PubMed=1469361;
RA   Abouzid A.M., Polston J.E., Hiebert E.;
RT   "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT   isolated from tomatoes in Florida.";
RL   J. Gen. Virol. 73:3225-3229(1992).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L14460; AAC32414.1; -.
```

```
DR   PIR; JQ1870; JQ1870.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (BY SIMILARITY).
SQ   SEQUENCE   361 AA;  40516 MW;  8138B65CEEAC6950 CRC64;


  Query Match             59.8%;  Score 214;  DB 1;  Length 361;
  Best Local Similarity   55.7%;  Pred. No. 1.8e-17;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:||||||||| |::||: | || | |  :|| ::||: |: :: | ||: |||
Db     110 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
VAL1_TYLCM
ID   VAL1_TYLCM     STANDARD;      PRT;   359 AA.
AC   P27260;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10833;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92107660; PubMed=1840676;
RA   Kheyr-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,
RA   Gronenborn B.;
RT   "Tomato yellow leaf curl virus from Sardinia is a
RT   whitefly-transmitted monopartite geminivirus.";
RL   Nucleic Acids Res. 19:6763-6769(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X61153; CAA43466.1; -.
DR   PIR; S22593; S22593.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  40733 MW;  9717B4A07C93EFA7 CRC64;


  Query Match             59.2%;  Score 212;  DB 1;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 3.1e-17;
  Matches   39;  Conservative   10;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| | |: | | : || :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT  12
VAL1_BCTV
ID   VAL1_BCTV      STANDARD;      PRT;   358 AA.
AC   P14991;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 protein (40.8 kDa protein).
OS   Beet curly top virus (BCTV).
OC   Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
OX   NCBI_TaxID=10840;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT   "The nucleotide sequence of an infectious clone of the geminivirus
RT   beet curly top virus.";
RL   EMBO J. 5:1761-1767(1986).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40889 MW;  39A45FE3C0B9C333 CRC64;


  Query Match             58.1%;  Score 208;  DB 1;  Length 358;
  Best Local Similarity   55.7%;  Pred. No. 9e-17;
  Matches   39;  Conservative   13;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| ||:||: | || |:|   ||||::|: |: |  | ||| :|
Db     110 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 169

Qy      61 DRIFDKTPEP 70
           :|| : |:|
Db     170 QKIFQRPPDP 179


RESULT  13
VAL1_TYLCV
ID   VAL1_TYLCV     STANDARD;      PRT;   357 AA.
AC   P27259;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10832;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92024070; PubMed=1926771;
RA   Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;
RT   "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RT   with a single genomic component.";
RL   Virology 185:151-161(1991).
```

```
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15656; CAA33688.1; -.
DR   PIR; D40779; QQCVC1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     219    226       ATP (POTENTIAL).
SQ   SEQUENCE   357 AA;  40678 MW;  939AB68E1AB3B2A7 CRC64;


  Query Match             53.1%;  Score 190;  DB 1;  Length 357;
  Best Local Similarity   60.7%;  Pred. No. 1.1e-14;
  Matches   37;  Conservative   11;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLDRI 63
           :|  |:||||||||| |::||| | || : | : || |::|| |: |: |||||:||||||
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171


RESULT  14
VAL1_SLCV
ID   VAL1_SLCV      STANDARD;      PRT;   347 AA.
AC   P29048;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   AL1 protein.
OS   Squash leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10829;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91082449; PubMed=1984668;
RA   Lazarowitz S.G., Lazdins I.B.;
RT   "Infectivity and complete nucleotide sequence of the cloned genomic
RT   components of a bipartite squash leaf curl geminivirus with a broad
RT   host range phenotype.";
RL   Virology 180:58-69(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M38183; AAC32410.1; ALT_INIT.
DR   PIR; C36785; QQCVS1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     218    225       ATP (POTENTIAL).
SQ   SEQUENCE   347 AA;  39110 MW;  AFDAEBDDE122110E CRC64;


  Query Match             32.1%;  Score 115;  DB 1;  Length 347;
  Best Local Similarity   36.4%;  Pred. No. 5.1e-06;
  Matches   24;  Conservative   14;  Mismatches   24;  Indels    4;  Gaps    1;

Qy       5 GEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLDRIF 64
           |:::| |    |     :|    |: | ||  || ||:   |: ::  :||| :|::|:|
Db     116 GQYKVSG----GSKSNKDDVYHNAVNAGSAGEALDIIKAGDPKTFIVNYHNLLANVERLF 171

Qy      65 DKTPEP 70
             | |||
Db     172 QKPPEP 177


RESULT  15
T2FA_DROME
ID   T2FA_DROME     STANDARD;      PRT;   577 AA.
AC   Q05913; Q9VI19;
DT   01-FEB-1994 (Rel. 28, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Transcription initiation factor IIF, alpha subunit (TFIIF-alpha)
DE   (Transcription factor 5, large chain) (TF5A).
GN   TFIIF-ALPHA OR CG10281.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93219133; PubMed=8464745;
RA   Gong D.-W., Horikoshi M., Nakatani Y.;
RT   "Analysis of cDNA encoding Drosophila transcription initiation factor
RT   TFIIF alpha (RAP74).";
RL   Nucleic Acids Res. 21:1492-1492(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93219101; PubMed=8464716;
RA   Kephart D., Price M.P., Burton Z.F., Finkelstein A.,
RA   Greenblalt J., Price D.H.;
RT   "Cloning of a Drosophila cDNA with sequence similarity to human
RT   transcription factor RAP74.";
RL   Nucleic Acids Res. 21:1319-1319(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
```

```
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [4]
RP   FUNCTION.
RX   MEDLINE=94230464; PubMed=8175788;
RA   Kephart D., Wang B.Q., Burton Z.F., Price D.H.;
RT   "Functional analysis of Drosophila factor 5 (TFIIF), a general
RT   transcription factor.";
RL   J. Biol. Chem. 269:13536-13543(1994).
CC   -!- FUNCTION: TFIIF IS A GENERAL TRANSCRIPTION INITIATION FACTOR THAT
CC       BINDS TO RNA POLYMERASE II AND HELPS TO RECRUIT IT TO THE
CC       INITIATION COMPLEX IN COLLABORATION WITH TFIIB. IT PROMOTES
CC       TRANSCRIPTION ELONGATION.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: TO OTHER TRANSCRIPTION FACTOR IIF, ALPHA SUBUNIT.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L10331; -; NOT_ANNOTATED_CDS.
DR   EMBL; X66982; CAA47391.1; -.
DR   EMBL; AE003675; AAF54125.1; -.
DR   PIR; S30237; S30237.
DR   TRANSFAC; T02173; -.
DR   FlyBase; FBgn0010282; TfIIF-alpha.
KW   Transcription regulation; DNA-binding; Nuclear protein.
FT   CONFLICT     20     20       A -> R (IN REF. 1).
FT   CONFLICT     76     76       E -> G (IN REF. 1).
FT   CONFLICT    528    528       P -> A (IN REF. 1).
SQ   SEQUENCE   577 AA;  64466 MW;  1E3DF88E3C833761 CRC64;


  Query Match             19.3%;  Score 69;  DB 1;  Length 577;
  Best Local Similarity   36.4%;  Pred. No. 1.9;
  Matches   20;  Conservative    9;  Mismatches   20;  Indels    6;  Gaps    1;

Qy      13 SARGGCQTSNDAAAAALAASSAAAALQIIRE------KIPEKYLFQFHNLNSNLD 61
           ||  |  ||  |||||  || :|::   ::|      |:|:|:     |   |:|
Db      10 SAASGSSTSAAAAAAASVASGSASSSANVQEFKIRVPKMPKKHHVMRFNATLNVD 64


Search completed: October 10, 2002, 02:09:51
Job time: 1349 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:48:42 ; Search time 69.63 Seconds
(without alignments)
96.600 Million cell updates/sec

Title: US-09-289-346A-3
Perfect score: 358
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	332	92.7	352	1	QQCVL1	AL1 protein - toma
2	243	67.9	361	1	QQCVPT	AL1 protein - pota
3	230	64.2	358	2	S07594	hypothetical prote
4	222	62.0	362	1	JQ1887	AL1 protein - toma
5	217	60.6	355	1	QQCVW1	AV1 protein - abut
6	216	60.3	349	2	JQ2300	replicase - pepper
7	216	60.3	349	2	S31875	AL1 protein - pepp
8	216	60.3	359	2	S39211	gene C1 protein -
9	214	59.8	358	1	JQ1870	AL1 protein - toma
10	213	59.5	351	2	JQ2327	AL1 protein - Indi
11	212	59.2	359	2	S22593	hypothetical prote
12	208	58.1	385	2	S28360	AL1 protein - beet
13	207	57.8	360	2	S59885	replication-associ
14	203	56.7	359	2	S39235	gene C1 protein -
15	190	53.1	357	1	QQCVC1	AL1 protein - toma
16	120	33.5	131	2	S45059	AC1 protein (clone
17	115	32.1	347	1	QQCVS1	AL1 protein - squa
18	69	19.3	577	2	S30237	transcription init
19	66	18.4	508	2	S59870	fork head domain p
20	66	18.4	840	2	T36175	probable large ATP
21	65	18.2	254	2	F97486	reco-like protein
22	65	18.2	254	2	AF2704	DNA repair protein
23	65	18.2	671	2	A35912	homeotic protein o
24	63	17.6	392	2	T44362	acetyl-CoA C-acety
25	62.5	17.5	700	2	T05841	spliceosome-associ
26	62.5	17.5	1077	2	A44067	serine-rich protei
27	62	17.3	409	2	T16781	hypothetical prote
28	61.5	17.2	317	2	D89961	conserved hypothet
29	61.5	17.2	327	2	F83408	hypothetical prote
30	61.5	17.2	705	2	T31157	hypothetical prote
31	61.5	17.2	1541	2	T02831	AAA protein L4171.
32	61	17.0	339	2	T26328	hypothetical prote
33	61	17.0	371	2	B69502	alcohol dehydrogen
34	61	17.0	1265	2	T47626	structural mainten
35	61	17.0	1510	2	T33100	hypothetical prote
36	60.5	16.9	509	2	T02942	O-succinylhomoseri
37	60.5	16.9	829	2	A34692	ecdysone-induced p
38	60.5	16.9	883	2	S04722	puff 74E protein -
39	60	16.8	314	2	JC5273	paired type homeob
40	59.5	16.6	201	2	A81380	probable flagellar
41	59.5	16.6	244	2	B86197	hypothetical prote
42	59.5	16.6	470	2	T46814	gamma-aminobutyrat
43	59.5	16.6	470	2	B95419	diaminobutyrate--p
44	59.5	16.6	481	2	A70091	probable phosphoes
45	59.5	16.6	509	2	T02940	O-succinylhomoseri

ALIGNMENTS

RESULT 1
QQCVL1
AL1 protein - tomato golden mosaic virus
C;Species: tomato golden mosaic virus
A;Note: host Nicotiana sp. (tobacco)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C;Accession: A04170
R;Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.
EMBO J. 3, 2197-2205, 1984
A;Title: Complete nucleotide sequence of the infectious cloned DNA components of toma
A;Reference number: A04163
A;Accession: A04170
A;Molecule type: DNA
A;Residues: 1-352 <HAM>
C;Comment: The genome consists of two circular, single-stranded DNA components, DNA A
.
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 92.7%; Score 332; DB 1; Length 352;
Best Local Similarity 92.9%; Pred. No. 4.8e-31;
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          |||||||||||||||||||||||||| || |||   ||||||||||||||||||||||||
Db    111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db    171 DRIFDKTPEP 180
```

RESULT 2
QQCVPT
AL1 protein - potato yellow mosaic virus (isolate Venezuela)
C;Species: potato yellow mosaic virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C;Accession: JU0364
R;Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.
J. Gen. Virol. 72, 1515-1520, 1991
A;Title: The nucleotide sequence of the infectious cloned DNA components of potato ye
A;Reference number: JU0362; MUID:91311403
A;Accession: JU0364
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-361 <COU>
A;Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             67.9%;  Score 243;  DB 1;  Length 361;
  Best Local Similarity   66.7%;  Pred. No. 1.2e-20;
  Matches   46;  Conservative   11;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:||||||||| || ||||| || : :  ||::||:||:|||:|||: ||
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 DRIFDKTPE 69
           |||| | ||
Db     170 DRIFMKAPE 178


RESULT   3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C;Species: cassava latent virus
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C;Accession: S07594
R;Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A;Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A;Reference number: S07590; MUID:90174930
A;Accession: S07594
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <MOR>
A;Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C;Genetics:
A;Map position: segment DNA1
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             64.2%;  Score 230;  DB 2;  Length 358;
  Best Local Similarity   60.0%;  Pred. No. 4.1e-19;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||| | || : | : || :||| :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   4
JQ1887
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1887
R;Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A;Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A;Reference number: JQ1885; MUID:93139778
A;Accession: JQ1887
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-362 <DRY>
A;Cross-references: GB:S53251
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             62.0%;  Score 222;  DB 1;  Length 362;
  Best Local Similarity   52.9%;  Pred. No. 3.5e-18;
  Matches   45;  Conservative    9;  Mismatches   15;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           || |||||:|||||||| |::||| | ||   | : || ::||  |: |: |||||||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   5
QQCVW1
AV1 protein - abutilon mosaic virus (isolate West India)
C;Species: abutilon mosaic virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Apr-1994
C;Accession: A36214
R;Frischmuth, T.; Zimmat, G.; Jeske, H.
Virology 178, 461-468, 1990
A;Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as
A;Reference number: A36214; MUID:91020984
A;Accession: A36214
A;Molecule type: DNA
A;Residues: 1-355 <FRI>
A;Cross-references: EMBL:X15983
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.6%;  Score 217;  DB 1;  Length 355;
  Best Local Similarity   58.6%;  Pred. No. 1.3e-17;
  Matches   41;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |  |||||:|||||||| ||:||: | || |    :|| |::|: |: |: | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT   6
JQ2300
replicase - pepper huasteco virus (component A)
N;Alternate names: ORF AL1 protein
C;Species: pepper huasteco virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C;Accession: JQ2300
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A;Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A;Reference number: JQ2299; MUID:94015007
A;Accession: JQ2300
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.3%;  Score 216;  DB 2;  Length 349;
  Best Local Similarity   58.6%;  Pred. No. 1.7e-17;
  Matches   41;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| |::||  | || ::||  |||||:|: |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   7
S31875
AL1 protein - pepper rizado amarillo virus
```

```
C;Species: pepper rizado amarillo virus
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Sep-1999
C;Accession: S31875
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera-Bu
submitted to the EMBL Data Library, February 1993
A;Description: Complete nucleotide sequence of pepper huasteco virus: analysis and compa
A;Reference number: S31872
A;Accession: S31875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
A;Note: the source is designated as pepper huasteco virus
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.3%;  Score 216;  DB 2;  Length 349;
  Best Local Similarity   58.6%;  Pred. No. 1.7e-17;
  Matches   41;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| |::||  | || ::||  |||||:|: |: :  |||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   8
S39211
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C;Accession: S39211
R;Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A;Description: High similarity among the tomato yellow leaf curl virus isolates from the
A;Reference number: S39209
A;Accession: S39211
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <NOR>
A;Cross-references: EMBL:Z25751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.3%;  Score 216;  DB 2;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 1.8e-17;
  Matches   40;  Conservative    9;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| | |: | | : || :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   9
JQ1870
AL1 protein - tomato mottle virus (isolate Florida)
C;Species: tomato mottle virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1870
R;Abouzid, A.M.; Polston, J.E.; Hiebert, E.
J. Gen. Virol. 73, 3225-3229, 1992
A;Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated from
A;Reference number: JQ1869; MUID:93107858
A;Accession: JQ1870
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <ABO>
A;Cross-references: GB:L14460
C;Genetics:
A;Map position: segment A
C;Superfamily: tomatc golden mosaic virus AL1 protein


  Query Match             59.8%;  Score 214;  DB 1;  Length 358;
  Best Local Similarity   55.7%;  Pred. No. 3e-17;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||: | || | |  :|| ::||: |: :: | ||: |||
Db     107 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 166

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     167 ERIFAKAPEP 176


RESULT  10
JQ2327
AL1 protein - Indian cassava mosaic virus
N;Alternate names: replication-associated protein
C;Species: Indian cassava mosaic virus
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C;Accession: JQ2327; S35883
R;Hong, Y.G.; Robinson, D.J.; Harrison, B.D.
J. Gen. Virol. 74, 2437-2443, 1993
A;Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-t
A;Reference number: JQ2326; MUID:94065670
A;Accession: JQ2327
A;Molecule type: DNA
A;Residues: 1-351 <HON>
A;Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.5%;  Score 213;  DB 2;  Length 351;
  Best Local Similarity   61.2%;  Pred. No. 3.8e-17;
  Matches   41;  Conservative   10;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLDRI 63
           || ||:|||||||| |::||| |||| : | : ||:|:||  |  ||  ||:::||||||
Db     113 WGTFQIDGRSARGGQQSANDAYAAALNSGSKSEALKILRELAPRDYLRDFHHISSNLDRI 172

Qy      64 FDKTPEP 70
           | | | |
Db     173 FTKPPPP 179


RESULT  11
S22593
hypothetical protein C4 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Sep-1999
C;Accession: S22593
R;Kheyr-Pour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn,
Nucleic Acids Res. 19, 6763-6769, 1991
A;Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monopa
A;Reference number: S22588; MUID:92107660
A;Accession: S22593
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-359 <KHE>
A;Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.2%;  Score 212;  DB 2;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 5.2e-17;
  Matches   39;  Conservative   10;  Mismatches   20;  Indels    0;  Gaps    0;
```

```
Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| | |: | | : || :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
             ::|   | |
Db     171 KVFQVPPAP 179
```

RESULT 12
S28360
AL1 protein - beet curly top virus
C;Species: beet curly top virus
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
C;Accession: S28360
R;Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
EMBO J. 5, 1761-1767, 1986
A;Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly to
A;Reference number: S28360
A;Accession: S28360
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-385 <STA>
A;Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             58.1%;  Score 208;  DB 2;  Length 385;
  Best Local Similarity   55.7%;  Pred. No. 1.6e-16;
  Matches   39;  Conservative   13;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:||||||||| ||:||: | || |:|   ||||::|: |: |  | ||| :|
Db     137 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 196

Qy      61 DRIFDKTPEP 70
            :|| : |:|
Db     197 QKIFQRPPDP 206
```

RESULT 13
S59885
replication-associated protein C1 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
C;Accession: S59885
R;Hong, Y.; Harrison, B.D.
submitted to the EMBL Data Library, February 1995
A;Description: Nucleotide sequences from tomato leaf curl viruses from different countri
d geminiviruses.
A;Reference number: S58346
A;Accession: S59885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <HON>
A;Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA88229.1; PID:g974211
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             57.8%;  Score 207;  DB 2;  Length 360;
  Best Local Similarity   60.6%;  Pred. No. 2e-16;
  Matches   40;  Conservative   12;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLDRI 63
           :| ||:|||||||| |::||| | |: : | |||| |:||| |: :: ||||||:|||||
Db     113 FGVFQIDGRSARGGQQSANDAYAEAINSGSKAAALDILREKAPKDFVLQFHNLNANLDRI 172

Qy      64 FDKTPE 69
           |  : |
Db     173 FTPSAE 178
```

RESULT 14
S39235
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Sep-1997
C;Accession: S39235
R;Crespi, S.; Noris, E.; Vaira, A.; Bosco, D.; Accotto, G.
submitted to the EMBL Data Library, December 1993
A;Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.
A;Reference number: S39233
A;Accession: S39235
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <CRE>
A;Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             56.7%;  Score 203;  DB 2;  Length 359;
  Best Local Similarity   53.6%;  Pred. No. 5.8e-16;
  Matches   37;  Conservative   12;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| | |: | | : || :|::  |  |:  |||::||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINARSKSEALDVIKQLAPRDYVLHFHNISSNLD 170

Qy      62 RIFDKTPEP 70
             ::|   | |
Db     171 KVFQVPPAP 179
```

RESULT 15
QQCVC1
AL1 protein - tomato yellow leaf curl virus
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C;Accession: D40779
R;Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.
Virology 185, 151-161, 1991
A;Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin
A;Reference number: A40779; MUID:92024070
A;Accession: D40779
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-357 <NAV>
A;Cross-references: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             53.1%;  Score 190;  DB 1;  Length 357;
  Best Local Similarity   60.7%;  Pred. No. 1.9e-14;
  Matches   37;  Conservative   11;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLDRI 63
           :|  |:|||||||| |::||| | || : | : || |::|| |: |: |||||:||||||
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171
```

Search completed: October 10, 2002, 01:48:43
Job time: 561 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:47:18 ; Search time 133.15 Seconds
(without alignments)
58.394 Million cell updates/sec

Title: US-09-289-346A-3
Perfect score: 358
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	358	100.0	70	21	AAB18679	Mutant peptide der
2	347	96.9	70	21	AAB18691	Mutant peptide der
3	343	95.8	70	21	AAB18690	Mutant peptide der
4	332	92.7	70	21	AAB18677	Peptide fragment f
5	332	92.7	356	21	AAB18687	Amino acid sequenc
6	324	90.5	70	21	AAB18685	Mutant peptide der
7	321	89.7	70	21	AAB18688	Mutant peptide der
8	320	89.4	70	21	AAB18692	Mutant peptide der
9	318	88.8	70	21	AAB18684	Mutant peptide der
10	317	88.5	70	21	AAB18678	Mutant peptide der
11	317	88.5	70	21	AAB18686	Mutant peptide der
12	316	88.3	70	21	AAB18689	Mutant peptide der
13	314	87.7	70	21	AAB18680	Mutant peptide der
14	311	86.9	70	21	AAB18681	Mutant peptide der
15	310	86.6	70	21	AAB18683	Mutant peptide der
16	308	86.0	70	21	AAB18682	Mutant peptide der
17	218	60.9	361	18	AAW34336	Tomato mottle viru
18	218	60.9	361	18	AAW34324	Tomato mottle viru
19	218	60.9	361	18	AAW34325	Tomato mottle viru
20	218	60.9	361	18	AAW34326	Tomato mottle viru
21	217	60.6	353	18	AAW34338	Bean golden mosaic
22	217	60.6	353	18	AAW34332	Bean golden mosaic
23	217	60.6	353	18	AAW34333	Bean golden mosaic
24	217	60.6	353	18	AAW34334	Bean golden mosaic
25	217	60.6	353	18	AAW34335	Bean golden mosaic
26	215	60.1	353	8	AAP70407	ORF 4 gene product
27	212	59.2	359	17	AAR88870	Sardinian tomato y
28	212	59.2	359	17	AAR88871	Sardinian tomato y
29	212	59.2	359	17	AAR88872	Sardinian tomato y
30	200.5	56.0	361	8	AAP70562	Product of ORF 4 f
31	198	55.3	362	19	AAW56495	Tobacco leaf curli
32	190	53.1	357	18	AAW34337	Tomato yellow leaf
33	190	53.1	357	18	AAW34329	Tomato yellow leaf
34	190	53.1	357	18	AAW34330	Tomato yellow leaf
35	190	53.1	357	18	AAW34331	Tomato yellow leaf
36	73.5	20.5	304	22	AAM39860	Human polypeptide
37	73.5	20.5	406	22	AAB92868	Human protein sequ
38	73.5	20.5	425	22	AAB74314	Human splicing fac
39	73.5	20.5	427	22	AAM41646	Human polypeptide
40	69	19.3	577	22	ABB64982	Drosophila melanog
41	66	18.4	508	22	ABB71345	Drosophila melanog
42	65	18.2	548	22	ABB59192	Drosophila melanog
43	63.5	17.7	100	22	ABG23455	Novel human diagno
44	63.5	17.7	606	22	ABG15588	Novel human diagno
45	63	17.6	956	22	ABB63221	Drosophila melanog

ALIGNMENTS

RESULT 1
AAB18679
ID AAB18679 standard; peptide; 70 AA.
XX
AC AAB18679;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 27
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 30
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 34
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 35
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 36
FT /note= "wild type residue replaced with Ala"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX

```
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 43; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match           100.0%;  Score 358;  DB 21;  Length 70;
  Best Local Similarity  100.0%;  Pred. No. 4.2e-39;
  Matches   70;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 tlvwgefqvdgrsarggcqtsndaaaaalaassaaaalqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   2
AAB18691
ID   AAB18691 standard; peptide; 70 AA.
XX
AC   AAB18691;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 34
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 35
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 36
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match           96.9%;  Score 347;  DB 21;  Length 70;
  Best Local Similarity  97.1%;  Pred. No. 1.1e-37;
  Matches   68;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          |||||||||||||||||||||||| || ||||||||||||||||||||||||||||||||
Db      1 tlvwgefqvdgrsarggcqtsndaaaealnassaaaalqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   3
AAB18690
ID   AAB18690 standard; peptide; 70 AA.
XX
AC   AAB18690;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 27
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 30
FT                   /note= "wild type residue replaced with Ala"
XX
```

PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DR WPI; 2000-618851/59.
XX
PT Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT -
XX
PS Disclosure; Page 49; 73pp; English.
XX
CC The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
SQ Sequence 70 AA;

Query Match 95.8%; Score 343; DB 21; Length 70;
Best Local Similarity 95.7%; Pred. No. 3.8e-37;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaaalaasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70
```

RESULT 4
AAB18677
ID AAB18677 standard; peptide; 70 AA.
XX
AC AAB18677;
XX
DT 22-JAN-2001 (first entry)
XX
DE Peptide fragment from Rep protein of TGMV (amino acids 110-179).
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Tomato golden mosaic virus.
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DR WPI; 2000-618851/59.
XX
PT Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT -
XX
PS Disclosure; Page 18; 73pp; English.
XX
CC The present sequence is derived from a geminivirus replication (Rep)
CC protein, which is also known as AL1. AL1 binds double-stranded DNA,
CC catalyses cleavage and ligation of single-stranded DNA, and interacts
CC with other viral and host proteins. Mutants of the AL1 protein are used
CC to produce transgenic plants. The mutation in AL1 is present in a
CC ribosome binding region, and expression of mutant AL1 protein imparts
CC increased resistance to geminivirus infection in the plant. Mutant AL1
CC proteins are useful for producing plants having increased resistance or
CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,
CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl
CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato
CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,
CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or
CC beet curly top virus.
XX
SQ Sequence 70 AA;

Query Match 92.7%; Score 332; DB 21; Length 70;
Best Local Similarity 92.9%; Pred. No. 1e-35;
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||||||||||||||||| || |||   ||||||||||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70
```

RESULT 5
AAB18687
ID AAB18687 standard; peptide; 356 AA.
XX
AC AAB18687;
XX
DT 22-JAN-2001 (first entry)
XX
DE Amino acid sequence of a geminivirus replication protein of TGMV.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 354
FT /note= "unspecified amino acid"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX

```
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 47-48; 73pp; English.
XX
CC   The present sequence represents a geminivirus replication (Rep)
CC   protein, which is also known as AL1. AL1 binds double-stranded DNA,
CC   catalyses cleavage and ligation of single-stranded DNA, and interacts
CC   with other viral and host proteins. Mutants of the AL1 protein are used
CC   to produce transgenic plants. The mutation in AL1 is present in a
CC   ribosome binding region, and expression of mutant AL1 protein imparts
CC   increased resistance to geminivirus infection in the plant. Mutant AL1
CC   proteins are useful for producing plants having increased resistance or
CC   reduced sensitivity to a geminivirus such as tomato golden mosaic virus,
CC   tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl
CC   virus, African cassava mosaic virus, Indian cassava mosaic virus, potato
CC   yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,
CC   squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or
CC   beet curly top virus.
XX
SQ   Sequence   356 AA;
```

```
  Query Match             92.7%;  Score 332;  DB 21;  Length 356;
  Best Local Similarity   92.9%;  Pred. No. 8e-35;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          ||||||||||||||||||||||||||| || |||   ||||||||||||||||||||||||
Db     110 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 169

Qy      61 DRIFDKTPEP 70
          ||||||||||
Db     170 drifdktpep 179
```

```
RESULT   6
AAB18685
ID   AAB18685 standard; peptide; 70 AA.
XX
AC   AAB18685;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 10
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
```

```
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 46; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;
```

```
  Query Match             90.5%;  Score 324;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 1.1e-34;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          |||||||| ||||||||||||||| || |||   ||||||||||||||||||||||||
Db       1 tlvwgefqvagrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
          ||||||||||
Db      61 drifdktpep 70
```

```
RESULT   7
AAB18688
ID   AAB18688 standard; peptide; 70 AA.
XX
AC   AAB18688;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 19
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 20
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
```

```
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 48; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;

  Query Match             89.7%;  Score 321;  DB 21;  Length 70;
  Best Local Similarity   90.0%;  Pred. No. 2.8e-34;
  Matches   63;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||||||||  |||||| || |||   ||||||||||||||||||||||||
Db       1 tlvwgefqvdgrsarggcaasndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT   8
AAB18692
ID   AAB18692 standard; peptide; 70 AA.
XX
AC   AAB18692;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 66
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 69
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 50; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;

  Query Match             89.4%;  Score 320;  DB 21;  Length 70;
  Best Local Similarity   90.0%;  Pred. No. 3.8e-34;
  Matches   63;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||||||||||||||| || |||   ||||||||||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||| || |
Db      61 drifdatpap 70


RESULT   9
AAB18684
ID   AAB18684 standard; peptide; 70 AA.
XX
AC   AAB18684;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 7
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 8
FT                   /note= "wild type residue replaced with Ala"
XX
```

```
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;
```

```
  Query Match             88.8%;  Score 318;  DB 21;  Length 70;
  Best Local Similarity   90.0%;  Pred. No. 6.9e-34;
  Matches   63;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||  |||||||||||||||||| || |||   |||||||||||||||||||||||||
Db       1 tlvwgeaavdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70
```

```
RESULT  10
AAB18678
ID   AAB18678 standard; peptide; 70 AA.
XX
AC   AAB18678;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 12
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 13
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 15
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 42-43; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;
```

```
  Query Match             88.5%;  Score 317;  DB 21;  Length 70;
  Best Local Similarity   88.6%;  Pred. No. 9.4e-34;
  Matches   62;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||  :| |||||||||||  || |||   |||||||||||||||||||||||||
Db       1 tlvwgefqvdgaaaaggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70
```

```
RESULT  11
AAB18686
ID   AAB18686 standard; peptide; 70 AA.
XX
AC   AAB18686;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
```

```
FH   Key             Location/Qualifiers
FT   Misc-difference 24
FT                   /note= "wild type residue replaced with Leu"
FT   Misc-difference 25
FT                   /note= "wild type residue replaced with Leu"
FT   Misc-difference 26
FT                   /note= "wild type residue replaced with Leu"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 46; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;
```

```
  Query Match             88.5%;  Score 317;  DB 21;  Length 70;
  Best Local Similarity   88.6%;  Pred. No. 9.4e-34;
  Matches   62;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||    || |||   ||||||||||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndlllealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70
```

```
RESULT  12
AAB18689
ID   AAB18689 standard; peptide; 70 AA.
XX
AC   AAB18689;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 22
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 23
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 48-49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;
```

```
  Query Match             88.3%;  Score 316;  DB 21;  Length 70;
  Best Local Similarity   90.0%;  Pred. No. 1.3e-33;
  Matches   63;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||  ||| || |||   ||||||||||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsaaaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70
```

```
RESULT  13
AAB18680
ID   AAB18680 standard; peptide; 70 AA.
XX
AC   AAB18680;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
```

```
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 42
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 43
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 44
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 43-44; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             87.7%;  Score 314;  DB 21;  Length 70;
  Best Local Similarity   88.6%;  Pred. No. 2.3e-33;
  Matches   62;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||||||||||||||||| || |||   |||||   ||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiiaaaipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  14
AAB18681
ID   AAB18681 standard; peptide; 70 AA.
XX
AC   AAB18681;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 47
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 48
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 49
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 44; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             86.9%;  Score 311;  DB 21;  Length 70;
  Best Local Similarity   88.6%;  Pred. No. 5.7e-33;
  Matches   62;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||||||||||||||||| || |||   ||||||||||   ||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipaaalfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  15
AAB18683
```

```
ID   AAB18683 standard; peptide; 70 AA.
XX
AC   AAB18683;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 59
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 61
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 62
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             86.6%;  Score 310;  DB 21;  Length 70;
  Best Local Similarity   88.6%;  Pred. No. 7.7e-33;
  Matches   62;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||||| || |||   ||||||||||||||||||||| |
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsal 60

Qy      61 DRIFDKTPEP 70
             ||||||||
Db      61 aaifdktpep 70
```

Search completed: October 10, 2002, 01:47:18
Job time: 527 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:01:59 ; Search time 118.38 Seconds
(without alignments)
102.295 Million cell updates/sec

Title: US-09-289-346A-4
Perfect score: 361
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	297	82.3	351	12	Q91R10	Q91r10 tomato seve
2	295	81.7	352	12	Q9E000	Q9e000 tomato rugo
3	281	77.8	226	12	O09727	O09727 leonurus mo
4	281	77.8	226	12	Q9WHF6	Q9whf6 tomato mild
5	279	77.3	361	12	Q67574	Q67574 bean golden
6	277	76.7	225	12	Q9QDB1	Q9qdb1 cowpea gold
7	274	75.9	314	12	Q9ELT8	Q9elt8 sweet potat
8	274	75.9	364	12	Q9QS55	Q9qs55 sweet potat
9	271	75.1	185	12	Q98693	Q98693 sida golden
10	266	73.7	149	12	P88975	P88975 macroptiliu
11	266	73.7	233	12	Q9YLA4	Q9yla4 macroptiliu
12	260	72.0	234	12	O39180	O39180 geminivirid
13	254	70.4	190	12	Q9Z089	Q9z089 tobacco lea
14	254	70.4	190	12	Q9Z084	Q9z084 tobacco lea
15	253	70.1	190	12	Q9W827	Q9w827 tobacco lea
16	253	70.1	208	12	Q9Z0C4	Q9z0c4 tobacco lea
17	253	70.1	359	12	Q91M88	Q91m88 tobacco lea
18	251	69.5	208	12	Q9Z0C0	Q9z0c0 tobacco lea
19	251	69.5	208	12	Q9Z0B8	Q9z0b8 tobacco lea
20	251	69.5	363	12	O72705	O72705 cotton leaf
21	251	69.5	363	12	O72719	O72719 cotton leaf
22	249	69.0	208	12	Q9Z0B6	Q9z0b6 tobacco lea
23	248	68.7	203	12	Q9Z083	Q9z083 tobacco lea
24	248	68.7	359	12	Q9YZV4	Q9yzv4 tomato yell
25	248	68.7	359	12	Q9YZV2	Q9yzv2 tomato yell
26	248	68.7	359	12	Q9YUX7	Q9yux7 tomato yell
27	248	68.7	359	12	Q9YL27	Q9yl27 tomato yell
28	248	68.7	363	12	O73577	O73577 cotton leaf
29	247	68.4	359	12	Q88942	Q88942 tomato yell
30	245	67.9	307	12	Q9IET7	Q9iet7 cotton leaf
31	245	67.9	361	12	O72723	O72723 cotton leaf
32	244	67.6	348	12	Q911W5	Q911w5 macroptiliu
33	243	67.3	359	12	Q91B86	Q91b86 ageratum ye
34	243	67.3	360	12	Q9DXI0	Q9dxi0 ageratum ye
35	243	67.3	362	12	Q9IN48	Q9in48 okra enatio
36	242	67.0	349	12	Q88888	Q88888 tomato pseu
37	242	67.0	362	12	O56816	O56816 chayote mos
38	241	66.8	208	12	Q9Z0A0	Q9z0a0 tobacco lea
39	241	66.8	354	12	Q9IN42	Q9in42 south afric
40	240	66.5	190	12	Q9Z086	Q9z086 tobacco lea
41	240	66.5	208	12	Q9Z0C6	Q9z0c6 tobacco lea
42	240	66.5	358	12	Q9JEA2	Q9jea2 cassava gem
43	239	66.2	363	12	O73494	O73494 okra yellow
44	237	65.7	306	12	Q9IET1	Q9iet1 cotton leaf
45	237	65.7	353	12	O72692	O72692 beet curly

ALIGNMENTS

RESULT 1
Q91R10
ID Q91R10 PRELIMINARY; PRT; 351 AA.
AC Q91R10;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE REP PROTEIN.
GN AC1.
OS Tomato severe rugose virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=158463;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MINAS GERAIS;
RA Rezende W.L., Goulart L.R., Parreira K.S., Figueiredo J.E.F.;
RT "The full-length DNA-A nucleotide sequence of a novel tomato-infecting
RT begomovirus, Tomato severe rugose virus, in Brazil.";
RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY029750; AAK50357.1; -.
SQ SEQUENCE 351 AA; 40122 MW; 87F937A4F873B6CF CRC64;

Query Match 82.3%; Score 297; DB 12; Length 351;
Best Local Similarity 78.6%; Pred. No. 2.9e-27;
Matches 55; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: ||||:||||||||||||:|||||||||| ||: ||||| :|||:|||||||||||
Db     111 TIEWGEFQIDGRSARGGCQTANDAAAEALNAPSKDVALQIIREKLPEKFLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| : |||
Db     171 DRIFARAPEP 180
```

RESULT 2
Q9E000
ID Q9E000 PRELIMINARY; PRT; 352 AA.

```
AC   Q9E000;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN.
GN   AC1.
OS   Tomato rugose mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=134599;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Fernandes J.J., Fontes E.P.B., Brommonschenkel S.H., Carvalho M.G.,
RA   Zambolim E.M., Zerbini F.M.;
RT   "Molecular Cloning and Characterization of Tomato rugose mosaic virus
RT   (TRMV), a Begomovirus Isolated from Tomatoes at Triangulo Mineiro,
RT   Minas Gerais, Brazil.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF291705; AAG15546.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   352 AA;  40012 MW;  47CD55838E24D613 CRC64;

  Query Match             81.7%;  Score 295;  DB 12;  Length 352;
  Best Local Similarity   78.6%;  Pred. No. 5.1e-27;
  Matches   55;  Conservative    7;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: |||||:||||||||||:|||||||||| ||: |||||   :|||:||||||||||||
Db     111 TIEWGEFQIDGRSARGGCQTANDAAAEALNAPSKDIALQIIREKMPEKFLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| : |||
Db     171 DRIFARAPEP 180


RESULT   3
O09727
ID   O09727      PRELIMINARY;      PRT;   226 AA.
AC   O09727;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   REP.
OS   Leonurus mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=58177;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LEMV- BRAZIL 1;
RA   Faria J.C., Maxwell D.P.;
RT   "Variability in geminivirus associated with Phaseolus vulgaris in
RT   Brazil.";
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U92532; AAB51157.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     226     226
SQ   SEQUENCE   226 AA;  25617 MW;  73CDB6E766083FC5 CRC64;

  Query Match             77.8%;  Score 281;  DB 12;  Length 226;
  Best Local Similarity   78.6%;  Pred. No. 1.4e-25;
  Matches   55;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: ||||||||||:||| || ||||||||||  |  |||||   :|||||||||||||||
Db     111 TVEWGEFQVDGRSSRGGQQTVNDAAAEALNAPDKRTALQIIKEKLPEKYLFQFHNLNSNL 170
```

```
Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180


RESULT   4
Q9WHF6
ID   Q9WHF6      PRELIMINARY;      PRT;   226 AA.
AC   Q9WHF6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   tomato mild mottle geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX   NCBI_TaxID=92943;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HN96-H5KW;
RA   Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA   Maxwell D.P.;
RT   "Molecular characterization and DNA-based detection methods for
RT   vegetable-infecting geminiviruses in Central America.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF131071; AAD33471.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     226     226
SQ   SEQUENCE   226 AA;  25941 MW;  2EA4116712871A23 CRC64;

  Query Match             77.8%;  Score 281;  DB 12;  Length 226;
  Best Local Similarity   74.3%;  Pred. No. 1.4e-25;
  Matches   52;  Conservative   11;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: |||||:||||||||| ||:|||||||||||||||||::||   :|:::|||:|||:|||
Db     111 TIEWGEFQIDGRSARGGQQTANDAAAEALNASSKEEAMRIIKEKLPKEFLFQYHNLSSNL 170

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180


RESULT   5
Q67574
ID   Q67574      PRELIMINARY;      PRT;   361 AA.
AC   Q67574;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE  REPLICATIVE PROTEIN.
GN   AL1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,
RA   Maxwell D.P., Russell D.R.;
RT   "Cloning of the complete DNA genomes of four bean-infecting
RT   geminiviruses and determining their infectivity by electric discharge
RT   particle acceleration.";
RL   Phytopathology 81:980-985(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,
RA   Morales F.J., Maxwell D.P.;
RT   "Differentiation of bean-infecting geminiviruses by nucleic acid
```

```
RT   hybridization probes and aspects of bean golden mosaic in Brazil.";
RL   Plant Dis. 75:336-342(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RT   "Genetic diversity in geminiviruses causing bean golden mosaic
RT   disease: The nucleotide sequence of the infectious cloned DNA
RT   components of a Brazilian isolate of bean golden mosaic geminivirus.";
RL   Submitted (MAY-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; M88686; AAA46312.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   361 AA;  41041 MW;  0094C7ACAF06B788 CRC64;


  Query Match             77.3%;  Score 279;  DB 12;  Length 361;
  Best Local Similarity   77.6%;  Pred. No. 4.3e-25;
  Matches   52;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLDRI 63
           || |||||||||| ||:||||:||||||||||||:|||   :|||:|||:|||:||||||
Db     113 WGHFQVDGRSARGGQQTANDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
           | | |:|
Db     173 FTKAPDP 179


RESULT   6
Q9QDB1
ID   Q9QDB1      PRELIMINARY;      PRT;   225 AA.
AC   Q9QDB1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   cowpea golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=69263;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CGMV-BR;
RA   Faria J.C.;
RT   "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT   Brazil.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF188708; AAF06318.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     225    225
SQ   SEQUENCE   225 AA;  25766 MW;  1089CB6BD8D15B5D CRC64;


  Query Match             76.7%;  Score 277;  DB 12;  Length 225;
  Best Local Similarity   77.6%;  Pred. No. 4.3e-25;
  Matches   52;  Conservative    7;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLDRI 63
           || ||:|||||||| || ||||:||||||||||:|||   :|||:|||:|||:||||||
Db     113 WGHFQIDGRSARGGQQTINDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
           | | |||
Db     173 FKKPPEP 179


RESULT   7
Q9ELT8
ID   Q9ELT8      PRELIMINARY;      PRT;   314 AA.
AC   Q9ELT8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION ASSOCIATION PROTEIN.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF288227; AAG01006.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   314 AA;  35153 MW;  6B6220613046943F CRC64;


  Query Match             75.9%;  Score 274;  DB 12;  Length 314;
  Best Local Similarity   79.4%;  Pred. No. 1.4e-24;
  Matches   54;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: ||||||||||||| ||:|||||||||| ||| |||||   :||||||||||| |||
Db     110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy      61 DRIFDKTP 68
           ||||   |
Db     170 DRIFSPPP 177


RESULT   8
Q9QS55
ID   Q9QS55      PRELIMINARY;      PRT;   364 AA.
AC   Q9QS55;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION INITIATION PROTEIN AC1.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A.;
RT   "Cloning of a DNA-A-like genomic component of sweet potato leaf curl
RT   virus : nucleotide sequence and phylogenetic relationships.";
RL   Online Publication.
DR   EMBL; AF104036; AAD47173.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   364 AA;  40680 MW;  5F79752431A09D6E CRC64;
```

```
  Query Match             75.9%;  Score 274;  DB 12;  Length 364;
  Best Local Similarity   79.4%;  Pred. No. 1.7e-24;
  Matches   54;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: |||||||||||||| ||:|||||||||| ||| |||||   :|||||||||| |||
Db     110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy      61 DRIFDKTP 68
           ||||   |
Db     170 DRIFSPPP 177


RESULT   9
Q98693
ID   Q98693      PRELIMINARY;      PRT;   185 AA.
AC   Q98693;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   AC1.
OS   sida golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51034;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA;
RA   Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RT   "Genetic Diversity among geminiviruses associated with the weed
RT   species Sida spp, Macroptilium lathyroides, and Wissadula amplissima
RT   from Jamaica.";
RL   Plant Dis. 81:1251-1258(1997).
DR   EMBL; U67926; AAB97865.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     185    185
SQ   SEQUENCE   185 AA;  20975 MW;  3913850A025A5EE1 CRC64;


  Query Match             75.1%;  Score 271;  DB 12;  Length 185;
  Best Local Similarity   70.0%;  Pred. No. 1.8e-24;
  Matches   49;  Conservative   12;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: || ||:|||||||| ||:|||||||||: :|||:||:||   :||:||||:|||:||:
Db      89 TIEWGVFQIDGRSARGGQQTANDAAAEALNSGTKEDALKIIREKLPERYLFQYHNLSSNI 148

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     149 DRIFSKPPEP 158


RESULT  10
P88975
ID   P88975      PRELIMINARY;      PRT;   149 AA.
AC   P88975;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   AC1.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICAN;
RA   Roye M.E.;
RL   Thesis (1996), Biochemistry, University of the West Indies, Jamaica.
DR   EMBL; U75278; AAB36919.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     149    149
SQ   SEQUENCE   149 AA;  16785 MW;  E4CF5EED4C9CD508 CRC64;


  Query Match             73.7%;  Score 266;  DB 12;  Length 149;
  Best Local Similarity   70.0%;  Pred. No. 5.4e-24;
  Matches   49;  Conservative   11;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: || ||:|||||||| ||||||||||||: :|| |::|:   :|||:|||:|||:|||
Db      52 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 111

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     112 DRIFMKDPEP 121


RESULT  11
Q9YLA4
ID   Q9YLA4      PRELIMINARY;      PRT;   233 AA.
AC   Q9YLA4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E.;
RT   "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
RT   from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E., McLaughlin W.A., Maxwell D.P.;
RT   "Molecular characterization of two distinct geminiviruses infecting M.
RT   lathyroides from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF098940; AAD17850.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     233    233
SQ   SEQUENCE   233 AA;  26356 MW;  AA490AF4D2166A02 CRC64;


  Query Match             73.7%;  Score 266;  DB 12;  Length 233;
  Best Local Similarity   70.0%;  Pred. No. 9.1e-24;
  Matches   49;  Conservative   11;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: || ||:|||||||| ||||||||||||: :|| |::|:   :|||:|||:|||:|||
Db     110 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 169

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     170 DRIFMKDPEP 179


RESULT  12
O39180
```

```
ID   O39180      PRELIMINARY;      PRT;   234 AA.
AC   O39180;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   AC1 (FRAGMENT).
OS   Geminiviridae.
OC   Viruses; ssDNA viruses.
OX   NCBI_TaxID=10811;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOMATO STRAIN;
RA   Guzman P., Arredondo C.R., Emmatty D., Portillo R.J., Gilbertson R.L.;
RT   "Partial Characterization of Two Whitefly-Transmitted Geminiviruses
RT   Infecting Tomatoes in Venezuela.";
RL   Plant Dis. 81:312-312(1997).
DR   EMBL; AF026553; AAB82605.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     234    234
SQ   SEQUENCE   234 AA;  26486 MW;  9ED8F0697105CD19 CRC64;


  Query Match             72.0%;  Score 260;  DB 12;  Length 234;
  Best Local Similarity   68.6%;  Pred. No. 4.8e-23;
  Matches   48;  Conservative   12;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: || ||:||||||||| ||:|||||||||: :||||::||   :|||:|||:|||:|:|
Db     110 TIEWGLFQIDGRSARGGQQTANDAAAEALNSGTKEEAMKIIKEKLPEKFLFQYHNLSSDL 169

Qy      61 DRIFDKTPEP 70
           |||| | | |
Db     170 DRIFMKAPNP 179


RESULT  13
Q9Z089
ID   Q9Z089      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z089;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA3-1, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001315; BAA34033.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21432 MW;  AAC093D1D1610FAD CRC64;


  Query Match             70.4%;  Score 254;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 1.9e-22;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           || || ||:|||||||||| :||| |||||||||| ||| ||   :|: ::|||:|||||||
Db      85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  14
Q9Z084
ID   Q9Z084      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z084;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA5-2, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA5;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001318; BAA34039.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  AAC1C2943E3F01AD CRC64;


  Query Match             70.4%;  Score 254;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 1.9e-22;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           || || ||:|||||||||| :||| |||||||||| ||| ||   :|: ::|||:|||||||
Db      85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  15
Q9W827
ID   Q9W827      PRELIMINARY;      PRT;   190 AA.
AC   Q9W827;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE GORA1-1, PARTIAL AND
DE   COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GORA1;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001303; BAA34010.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
```

```
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  93C3742A8EBDB7EB CRC64;


  Query Match             70.1%;  Score 253;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 2.6e-22;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           || || |||||||||||||| :||| |||||||:|| ||| ||   :|: ::||:|||||||
Db      85 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFSNSSFDQVPE 169


Search completed: October 10, 2002, 02:01:59
Job time: 1092 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:09:51 ; Search time 34.62 Seconds
(without alignments)
78.289 Million cell updates/sec

Title: US-09-289-346A-4
Perfect score: 361
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	346	95.8	352	1	VAL1_TGMV	P03567 tomato gold
2	249	69.0	361	1	VAL1_PYMVV	P27258 potato yell
3	240	66.5	358	1	VAL1_CLVK	P14982 cassava lat
4	240	66.5	358	1	VAL1_CLVN	P14972 cassava lat
5	233	64.5	362	1	VAL1_TYLCA	P36279 tomato yell
6	229	63.4	359	1	VAL1_TYLCU	P38609 tomato yell
7	222	61.5	349	1	VAL1_PHUV	Q06923 pepper huas
8	222	61.5	359	1	VAL1_TYLCM	P27260 tomato yell
9	217	60.1	353	1	VAL1_BGMV	P05175 bean golden
10	216	59.8	355	1	VAL1_ABMVW	P21947 abutilon mo
11	213	59.0	358	1	VAL1_BCTV	P14991 beet curly
12	210	58.2	361	1	VAL1_TMOV	Q06657 tomato mott
13	200	55.4	357	1	VAL1_TYLCV	P27259 tomato yell
14	128	35.5	347	1	VAL1_SLCV	P29048 squash leaf
15	61.5	17.0	299	1	Y175_HELPJ	Q9zmq7 helicobacte
16	61.5	17.0	334	1	G3P_BACST	P00362 bacillus st
17	61.5	17.0	335	1	G3P_BACCO	P15115 bacillus co
18	61.5	17.0	1044	1	BUB1_SCHPO	O94751 schizosacch
19	59	16.3	1483	1	CYP1_YEAST	P12351 saccharomyc
20	58.5	16.2	297	1	RRPP_RABVP	P06747 rabies viru
21	58	16.1	492	1	MOT3_MOUSE	O35308 mus musculu
22	58	16.1	492	1	MOT3_RAT	O70461 rattus norv
23	58	16.1	555	1	GLPD_BACSU	P18158 bacillus su
24	58	16.1	601	1	CYSJ_BUCAI	P57503 buchnera ap
25	57	15.8	223	1	DGK2_LACAC	Q59484 lactobacill
26	57	15.8	316	1	SPDE_COFAR	O82147 coffea arab
27	57	15.8	379	1	CYB_ASTPE	Q33818 asterina pe
28	57	15.8	387	1	Y4PF_RHISN	P55615 rhizobium s
29	57	15.8	836	1	GCSR_HUMAN	Q99062 homo sapien
30	57	15.8	1287	1	SKI2_YEAST	P35207 saccharomyc
31	56.5	15.7	863	1	AMPN_CAUCR	P37893 caulobacter
32	56	15.5	367	1	LHX4_MOUSE	P53776 mus musculu
33	56	15.5	577	1	PTLB_LACCA	P24400 lactobacill
34	56	15.5	874	1	SLAP_BACLI	P49052 bacillus li
35	55.5	15.4	299	1	Y175_HELPY	P56112 helicobacte
36	55.5	15.4	323	1	VAL1_PASVK	Q00338 panicum str
37	55.5	15.4	397	1	METL_RAT	P13444 rattus norv
38	55.5	15.4	428	1	GBA1_CANAL	P28868 candida alb
39	55	15.2	355	1	CRTB_RHOSH	P54905 rhodobacter
40	55	15.2	524	1	MPPA_MOUSE	Q9dc61 mus musculu
41	55	15.2	524	1	MPPA_RAT	P20069 rattus norv
42	55	15.2	525	1	MPPA_HUMAN	Q10713 homo sapien
43	54.5	15.1	297	1	RRPP_RABVE	P22559 rabies viru
44	54.5	15.1	405	1	WCAC_ECOLI	P71237 escherichia
45	54.5	15.1	445	1	ACSC_MOOTH	Q07340 moorella th

ALIGNMENTS

RESULT 1
VAL1_TGMV
ID VAL1_TGMV STANDARD; PRT; 352 AA.
AC P03567;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein.
GN AC1.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10831;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;
RT "Complete nucleotide sequence of the infectious cloned DNA components
RT of tomato golden mosaic virus: potential coding regions and regulatory
RT sequences.";
RL EMBO J. 3:2197-2205(1984).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; K02029; -; NOT_ANNOTATED_CDS.
DR PIR; A04170; QQCVL1.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR ProDom; PD000736; Gemini_AL1; 1.
KW ATP-binding.
FT NP_BIND 223 230 ATP (POTENTIAL).
SQ SEQUENCE 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;

Query Match 95.8%; Score 346; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 5.2e-34;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |||||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
VAL1_PYMVV
ID VAL1_PYMVV STANDARD; PRT; 361 AA.

```
AC   P27258;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 protein.
OS   Potato yellow mosaic virus (isolate Venezuela).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10828;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91311403; PubMed=1856690;
RA   Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT   "The nucleotide sequence of the infectious cloned DNA components of
RT   potato yellow mosaic virus.";
RL   J. Gen. Virol. 72:1515-1520(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00940; BAA00782.1; -.
DR   PIR; JU0364; QQCVPT.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   361 AA;  40850 MW;  5627A33BF1264383 CRC64;


  Query Match             69.0%;  Score 249;  DB 1;  Length 361;
  Best Local Similarity   68.1%;  Pred. No. 2.2e-22;
  Matches   47;  Conservative   10;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: || ||:|||||||| || ||||||||: :|| |::||   :|||:|||:|||: ||
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 DRIFDKTPE 69
           |||| | ||
Db     170 DRIFMKAPE 178


RESULT   3
VAL1_CLVK
ID   VAL1_CLVK      STANDARD;      PRT;   358 AA.
AC   P14982;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (40.4 kDa protein).
GN   AC1.
OS   Cassava latent virus (strain West Kenyan 844).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10818;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Gay M.R.;
RT   "Nucleotide sequence of cassava latent virus DNA.";
RL   Nature 301:260-262(1983).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J02057; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40346 MW;  ED173E753EE92D69 CRC64;


  Query Match             66.5%;  Score 240;  DB 1;  Length 358;
  Best Local Similarity   61.4%;  Pred. No. 2.7e-21;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||| |:|||: || ||| :|   :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   4
VAL1_CLVN
ID   VAL1_CLVN      STANDARD;      PRT;   358 AA.
AC   P14972;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (40.4 kDa protein).
GN   AC1.
OS   Cassava latent virus (strain Nigerian).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10819;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90174930; PubMed=2308831;
RA   Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT   "Nucleotide sequence of the infectious cloned DNA components of
RT   African cassava mosaic virus (Nigerian strain).";
RL   Nucleic Acids Res. 18:197-198(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X17095; CAA34953.1; -.
DR   PIR; S07594; S07594.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40435 MW;  1DB16BB0CB2D5E2C CRC64;


  Query Match             66.5%;  Score 240;  DB 1;  Length 358;
  Best Local Similarity   61.4%;  Pred. No. 2.7e-21;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||| |:|||: || ||| :|   :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168
```

```
Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   5
VAL1_TYLCA
ID   VAL1_TYLCA     STANDARD;      PRT;   362 AA.
AC   P36279;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=36447;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93139778; PubMed=8423446;
RA   Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;
RT   "Nucleotide sequence and genome organization of tomato leaf curl
RT   geminivirus.";
RL   J. Gen. Virol. 74:147-151(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
DR   PIR; JQ1887; JQ1887.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   362 AA;  41197 MW;  343E7184B4704098 CRC64;


  Query Match             64.5%;  Score 233;  DB 1;  Length 362;
  Best Local Similarity   54.1%;  Pred. No. 1.9e-20;
  Matches   46;  Conservative    9;  Mismatches   14;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           || ||||||:|||||||| |::||| |:||| || ||| ::    |: |: |||||||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   6
VAL1_TYLCU
ID   VAL1_TYLCU     STANDARD;      PRT;   359 AA.
AC   P38609;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=37139;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94256836; PubMed=8198442;
RA   Noris E., Hidalgo E., Accotto G., Moriones E.;
RT   "High similarity among the tomato yellow leaf curl virus isolates
RT   from the west Mediterranean basin: the nucleotide sequence of an
RT   infectious clone from Spain.";
RL   Arch. Virol. 135:165-170(1994).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z25751; CAA81026.1; -.
DR   PIR; S39211; S39211.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  41065 MW;  2D170A51EF80A3EC CRC64;


  Query Match             63.4%;  Score 229;  DB 1;  Length 359;
  Best Local Similarity   60.9%;  Pred. No. 5.5e-20;
  Matches   42;  Conservative    8;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLD 61
           | || ||:||||||||| ||:||| |:|:|| || ||| :|    |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   7
VAL1_PHUV
ID   VAL1_PHUV      STANDARD;      PRT;   349 AA.
AC   Q06923;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AL1.
OS   Pepper huasteco virus (PHV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=28349;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94015007; PubMed=8409944;
RA   Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA   Rivera-Bustamante R.F.;
RT   "Complete nucleotide sequence of pepper huasteco virus: analysis and
RT   comparison with bipartite geminiviruses.";
RL   J. Gen. Virol. 74:2225-2231(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70418; CAA49856.1; -.
DR   PIR; S31875; S31875.
DR   PIR; JQ2300; JQ2300.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (BY SIMILARITY).
SQ   SEQUENCE   349 AA;  39722 MW;  D5F4E76CD56370F4 CRC64;


  Query Match             61.5%;  Score 222;  DB 1;  Length 349;
```

```
  Best Local Similarity   60.0%;  Pred. No. 3.7e-19;
  Matches   42;  Conservative   11;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| |::|| |:|||::| |||||||    |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   8
VAL1_TYLCM
ID   VAL1_TYLCM     STANDARD;      PRT;   359 AA.
AC   P27260;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10833;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92107660; PubMed=1840676;
RA   Kheyr-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,
RA   Gronenborn B.;
RT   "Tomato yellow leaf curl virus from Sardinia is a
RT   whitefly-transmitted monopartite geminivirus.";
RL   Nucleic Acids Res. 19:6763-6769(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X61153; CAA43466.1; -.
DR   PIR; S22593; S22593.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  40733 MW;  9717B4A07C93EFA7 CRC64;


  Query Match             61.5%;  Score 222;  DB 1;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 3.8e-19;
  Matches   40;  Conservative   10;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|    |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   9
VAL1_BGMV
ID   VAL1_BGMV      STANDARD;      PRT;   353 AA.
AC   P05175;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   AL1 protein (40.2 kDa protein).
GN   AC1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Howarth A.J., Caton J., Bossert M., Goodman R.M.;
RT   "Nucleotide sequence of bean golden mosaic virus and a model for gene
RT   regulation in geminiviruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M10070; AAA46318.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   353 AA;  40190 MW;  80FA779DF6029A34 CRC64;


  Query Match             60.1%;  Score 217;  DB 1;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 1.5e-18;
  Matches   42;  Conservative   11;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: ||:|||||||||| |::||: |:|||| | | || |:    |: |: | ||: |||
Db     110 TIEWGQFQVDGRSARGGQQSANDSYAKALNADSIESALTILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFFKVPEP 179


RESULT  10
VAL1_ABMVW
ID   VAL1_ABMVW     STANDARD;      PRT;   355 AA.
AC   P21947;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AC1.
OS   Abutilon mosaic virus (isolate West India).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10816;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91020984; PubMed=2219703;
RA   Frischmuth T., Zimmat G., Jeske H.;
RT   "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT   as well as eukaryotic features.";
RL   Virology 178:461-468(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
```

```
DR   EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR   PIR; A36214; QQCVW1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   355 AA;  40257 MW;  16A2CA8A63251E95 CRC64;


  Query Match             59.8%;  Score 216;  DB 1;  Length 355;
  Best Local Similarity   58.6%;  Pred. No. 2e-18;
  Matches   41;  Conservative   10;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |  |||||:||||||||| ||:||: |:||||   : || |:    |: |: | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
VAL1_BCTV
ID   VAL1_BCTV      STANDARD;      PRT;   358 AA.
AC   P14991;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 protein (40.8 kDa protein).
OS   Beet curly top virus (BCTV).
OC   Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
OX   NCBI_TaxID=10840;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT   "The nucleotide sequence of an infectious clone of the geminivirus
RT   beet curly top virus.";
RL   EMBO J. 5:1761-1767(1986).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40889 MW;  39A45FE3C0B9C333 CRC64;


  Query Match             59.0%;  Score 213;  DB 1;  Length 358;
  Best Local Similarity   55.7%;  Pred. No. 4.6e-18;
  Matches   39;  Conservative   14;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| ||:||: |:||||:| ::||||:    |: |  | ||| :|
Db     110 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 169

Qy      61 DRIFDKTPEP 70
             :|| : |:|
Db     170 QKIFQRPPDP 179


RESULT  12
VAL1_TMOV
ID   VAL1_TMOV      STANDARD;      PRT;   361 AA.
AC   Q06657;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AL1.
OS   Tomato mottle virus (isolate Florida) (TMoV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=36449;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93107858; PubMed=1469361;
RA   Abouzid A.M., Polston J.E., Hiebert E.;
RT   "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT   isolated from tomatoes in Florida.";
RL   J. Gen. Virol. 73:3225-3229(1992).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L14460; AAC32414.1; -.
DR   PIR; JQ1870; JQ1870.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (BY SIMILARITY).
SQ   SEQUENCE   361 AA;  40516 MW;  8138B65CEEAC6950 CRC64;


  Query Match             58.2%;  Score 210;  DB 1;  Length 361;
  Best Local Similarity   54.3%;  Pred. No. 1.1e-17;
  Matches   38;  Conservative   15;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||: |:|||| | : || ::    |: :: | ||: |||
Db     110 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  13
VAL1_TYLCV
ID   VAL1_TYLCV     STANDARD;      PRT;   357 AA.
AC   P27259;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10832;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92024070; PubMed=1926771;
RA   Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;
RT   "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RT   with a single genomic component.";
RL   Virology 185:151-161(1991).
```

```
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15656; CAA33688.1; -.
DR   PIR; D40779; QQCVC1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     219    226       ATP (POTENTIAL).
SQ   SEQUENCE   357 AA;  40678 MW;  939AB68E1AB3B2A7 CRC64;


  Query Match             55.4%;  Score 200;  DB 1;  Length 357;
  Best Local Similarity   63.9%;  Pred. No. 1.6e-16;
  Matches   39;  Conservative    9;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLDRI 63
           :|  |:||||||| |::||| |||||: || ||| |:    |: |: |||||:||||||
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171


RESULT  14
VAL1_SLCV
ID   VAL1_SLCV      STANDARD;      PRT;   347 AA.
AC   P29048;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   AL1 protein.
OS   Squash leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10829;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91082449; PubMed=1984668;
RA   Lazarowitz S.G., Lazdins I.B.;
RT   "Infectivity and complete nucleotide sequence of the cloned genomic
RT   components of a bipartite squash leaf curl geminivirus with a broad
RT   host range phenotype.";
RL   Virology 180:58-69(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M38183; AAC32410.1; ALT_INIT.
DR   PIR; C36785; QQCVS1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     218    225       ATP (POTENTIAL).
SQ   SEQUENCE   347 AA;  39110 MW;  AFDAEBDDE122110E CRC64;


  Query Match             35.5%;  Score 128;  DB 1;  Length 347;
  Best Local Similarity   39.4%;  Pred. No. 6.7e-08;
  Matches   26;  Conservative   13;  Mismatches   23;  Indels    4;  Gaps    1;

Qy       5 GEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLDRIF 64
           |:::| |    |     :|    |:|| |  ||| || |  |: ::  :||| :|::|:|
Db     116 GQYKVSG----GSKSNKDDVYHNAVNAGSAGEALDIIKAGDPKTFIVNYHNLLANVERLF 171

Qy      65 DKTPEP 70
             | |||
Db     172 QKPPEP 177


RESULT  15
Y175_HELPJ
ID   Y175_HELPJ     STANDARD;      PRT;   299 AA.
AC   Q9ZMQ7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein JHP0161 precursor.
GN   JHP0161.
OS   Helicobacter pylori J99 (Campylobacter pylori J99).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
OX   NCBI_TaxID=85963;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99120557; PubMed=9923682;
RA   Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA   Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA   Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA   Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA   Trust T.J.;
RT   "Genomic sequence comparison of two unrelated isolates of the human
RT   gastric pathogen Helicobacter pylori.";
RL   Nature 397:176-180(1999).
CC   -!- SIMILARITY: BELONGS TO THE PPIC/PARVULIN FAMILY OF ROTAMASES.
CC       STRONG, TO C.JEJUNI CBF2.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001454; AAD05744.1; -.
DR   HSSP; Q13526; 1PIN.
DR   InterPro; IPR000297; Rotamase.
DR   Pfam; PF00639; Rotamase; 1.
DR   PROSITE; PS01096; PPIC_PPIASE_1; 1.
DR   PROSITE; PS50198; PPIC_PPIASE_2; 1.
KW   Hypothetical protein; Isomerase; Rotamase; Signal; Complete proteome.
FT   SIGNAL       1     21       POTENTIAL.
FT   CHAIN       22    299       HYPOTHETICAL PROTEIN JHP0161.
FT   DOMAIN     154    253       PPIC.
SQ   SEQUENCE   299 AA;  34040 MW;  9C037B1CD1110143 CRC64;


  Query Match             17.0%;  Score 61.5;  DB 1;  Length 299;
  Best Local Similarity   28.1%;  Pred. No. 5.3;
  Matches   18;  Conservative   12;  Mismatches   13;  Indels   21;  Gaps    2;

Qy      22 NDAAAEALNAS------------------SKEEALQIIAAAIPEKYLFQFHNLNSNLDR 62
           |:| || || :                   :|::| ::    |||| :  |:  |:| |:
Db      93 NEAKAEKLNQTPEFKAMMEAVKKQALVEFWAKKQAEEVKKIQIPEKEMQDFY--NANKDQ 150

Qy      63 IFDK 66
           :| |
Db     151 LFVK 154
```

Search completed: October 10, 2002, 02:09:51
Job time: 1349 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:48:43 ; Search time 69.63 Seconds
(without alignments)
96.600 Million cell updates/sec

Title: US-09-289-346A-4
Perfect score: 361
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	346	95.8	352	1	QQCVL1	AL1 protein - toma
2	249	69.0	361	1	QQCVPT	AL1 protein - pota
3	240	66.5	358	2	S07594	hypothetical prote
4	233	64.5	362	1	JQ1887	AL1 protein - toma
5	229	63.4	359	2	S39211	gene C1 protein -
6	222	61.5	349	2	JQ2300	replicase - pepper
7	222	61.5	349	2	S31875	AL1 protein - pepp
8	222	61.5	359	2	S22593	hypothetical prote
9	219	60.7	359	2	S39235	gene C1 protein -
10	216	59.8	351	2	JQ2327	AL1 protein - Indi
11	216	59.8	355	1	QQCVW1	AV1 protein - abut
12	213	59.0	385	2	S28360	AL1 protein - beet
13	210	58.2	358	1	JQ1870	AL1 protein - toma
14	200	55.4	357	1	QQCVC1	AL1 protein - toma
15	199	55.1	360	2	S59885	replication-associ
16	134	37.1	131	2	S45059	AC1 protein (clone
17	128	35.5	347	1	QQCVS1	AL1 protein - squa
18	72	19.9	587	2	JC1419	Fc gamma (IgG) rec
19	68	18.8	714	2	C95382	probable ferrichro
20	63.5	17.6	584	2	T19061	hypothetical prote
21	62.5	17.3	1006	2	S74992	hypothetical prote
22	61.5	17.0	299	2	B71967	probable peptidyl-
23	61.5	17.0	335	1	DEBSGF	glyceraldehyde-3-p
24	61.5	17.0	481	2	A70091	probable phosphoes
25	61.5	17.0	1044	2	T43800	protein kinase bub
26	61.5	17.0	1203	2	T17415	mycelial surface a
27	61.5	17.0	1229	2	T48959	kinesin-like prote
28	60.5	16.8	159	2	C71838	NADH dehydrogenase
29	60.5	16.8	159	2	E64677	NADH dehydrogenase
30	60.5	16.8	819	2	AI3197	TonB-dependent rec
31	60	16.6	767	2	F71479	hypothetical prote
32	60	16.6	1008	2	T41244	SEC14 protein homo
33	59	16.3	160	2	G82060	hypothetical prote
34	59	16.3	201	2	G97513	NADH dehydrogenase
35	59	16.3	201	2	AE2732	NADH ubiqionone ox
36	59	16.3	228	2	F90130	hypothetical prote
37	59	16.3	338	2	AG3582	iron(III)-binding
38	59	16.3	1502	1	RGBYH1	CYC1/CYP3 transcri
39	58.5	16.2	297	1	MNVNRV	nonstructural prot
40	58.5	16.2	357	2	T02246	hypothetical prote
41	58.5	16.2	569	2	H87926	protein C17D12.7 [
42	58.5	16.2	642	2	D90558	topoisomerase iv s
43	58.5	16.2	769	2	F81742	conserved hypothet
44	58.5	16.2	2135	2	T14602	variant-specific s
45	58	16.1	345	2	AD3024	hypothetical prote

ALIGNMENTS

RESULT 1
QQCVL1
AL1 protein - tomato golden mosaic virus
C;Species: tomato golden mosaic virus
A;Note: host Nicotiana sp. (tobacco)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C;Accession: A04170
R;Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.
EMBO J. 3, 2197-2205, 1984
A;Title: Complete nucleotide sequence of the infectious cloned DNA components of toma
A;Reference number: A04163
A;Accession: A04170
A;Molecule type: DNA
A;Residues: 1-352 <HAM>
C;Comment: The genome consists of two circular, single-stranded DNA components, DNA A
.
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 95.8%; Score 346; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 4e-33;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |||||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
QQCVPT
AL1 protein - potato yellow mosaic virus (isolate Venezuela)
C;Species: potato yellow mosaic virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C;Accession: JU0364
R;Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.
J. Gen. Virol. 72, 1515-1520, 1991
A;Title: The nucleotide sequence of the infectious cloned DNA components of potato ye
A;Reference number: JU0362; MUID:91311403
A;Accession: JU0364
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-361 <COU>
A;Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             69.0%;  Score 249;  DB 1;  Length 361;
  Best Local Similarity   68.1%;  Pred. No. 1.2e-21;
  Matches   47;  Conservative   10;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: || ||:|||||||| || |||||||||: :|| |::||   :|||:|||:|||: ||
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 DRIFDKTPE 69
           |||| | ||
Db     170 DRIFMKAPE 178


RESULT   3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C;Species: cassava latent virus
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C;Accession: S07594
R;Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A;Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A;Reference number: S07590; MUID:90174930
A;Accession: S07594
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <MOR>
A;Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C;Genetics:
A;Map position: segment DNA1
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             66.5%;  Score 240;  DB 2;  Length 358;
  Best Local Similarity   61.4%;  Pred. No. 1.4e-20;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||| |:|||: || ||| :|   :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   4
JQ1887
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1887
R;Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A;Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A;Reference number: JQ1885; MUID:93139778
A;Accession: JQ1887
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-362 <DRY>
A;Cross-references: GB:S53251
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             64.5%;  Score 233;  DB 1;  Length 362;
  Best Local Similarity   54.1%;  Pred. No. 9.8e-20;
  Matches   46;  Conservative    9;  Mismatches   14;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           || |||||:|||||||| |::||| |:|||  || ||| ::    |: |: |||||||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   5
S39211
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C;Accession: S39211
R;Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A;Description: High similarity among the tomato yellow leaf curl virus isolates from
A;Reference number: S39209
A;Accession: S39211
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <NOR>
A;Cross-references: EMBL:Z25751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             63.4%;  Score 229;  DB 2;  Length 359;
  Best Local Similarity   60.9%;  Pred. No. 2.9e-19;
  Matches   42;  Conservative    8;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || ||| :|    |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   6
JQ2300
replicase - pepper huasteco virus (component A)
N;Alternate names: ORF AL1 protein
C;Species: pepper huasteco virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C;Accession: JQ2300
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A;Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A;Reference number: JQ2299; MUID:94015007
A;Accession: JQ2300
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             61.5%;  Score 222;  DB 2;  Length 349;
  Best Local Similarity   60.0%;  Pred. No. 1.9e-18;
  Matches   42;  Conservative   11;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| |::||  |:|||::| |||||||    |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   7
S31875
AL1 protein - pepper rizado amarillo virus
C;Species: pepper rizado amarillo virus
```

```
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Sep-1999
C;Accession: S31875
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera-Bu
submitted to the EMBL Data Library, February 1993
A;Description: Complete nucleotide sequence of pepper huasteco virus: analysis and compa
A;Reference number: S31872
A;Accession: S31875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
A;Note: the source is designated as pepper huasteco virus
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             61.5%;  Score 222;  DB 2;  Length 349;
  Best Local Similarity   60.0%;  Pred. No. 1.9e-18;
  Matches   42;  Conservative   11;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: ||||:|||||||| |::|| |:|||::| ||||||| |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   8
S22593
hypothetical protein C4 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Sep-1999
C;Accession: S22593
R;Kheyr-Pour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.
Nucleic Acids Res. 19, 6763-6769, 1991
A;Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monoparti
A;Reference number: S22588; MUID:92107660
A;Accession: S22593
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-359 <KHE>
A;Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             61.5%;  Score 222;  DB 2;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 1.9e-18;
  Matches   40;  Conservative   10;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|    |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   9
S39235
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Sep-1997
C;Accession: S39235
R;Crespi, S.; Noris, E.; Vaira, A.; Bosco, D.; Accotto, G.
submitted to the EMBL Data Library, December 1993
A;Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.
A;Reference number: S39233
A;Accession: S39235
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <CRE>
A;Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.7%;  Score 219;  DB 2;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 4.4e-18;
  Matches   40;  Conservative   10;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || ||| :|    |  |:  |||::||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINARSKSEALDVIKQLAPRDYVLHFHNISSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT  10
JQ2327
AL1 protein - Indian cassava mosaic virus
N;Alternate names: replication-associated protein
C;Species: Indian cassava mosaic virus
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C;Accession: JQ2327; S35883
R;Hong, Y.G.; Robinson, D.J.; Harrison, B.D.
J. Gen. Virol. 74, 2437-2443, 1993
A;Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-t
A;Reference number: JQ2326; MUID:94065670
A;Accession: JQ2327
A;Molecule type: DNA
A;Residues: 1-351 <HON>
A;Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.8%;  Score 216;  DB 2;  Length 351;
  Best Local Similarity   61.2%;  Pred. No. 9.7e-18;
  Matches   41;  Conservative    9;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLDRI 63
           || ||:|||||||| |::||| | |||: || |||:|:    |  ||  ||:::||||||
Db     113 WGTFQIDGRSARGGQQSANDAYAAALNSGSKSEALKILRELAPRDYLRDFHHISSNLDRI 172

Qy      64 FDKTPEP 70
           | | | |
Db     173 FTKPPPP 179


RESULT  11
QQCVW1
AV1 protein - abutilon mosaic virus (isolate West India)
C;Species: abutilon mosaic virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Apr-1994
C;Accession: A36214
R;Frischmuth, T.; Zimmat, G.; Jeske, H.
Virology 178, 461-468, 1990
A;Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as
A;Reference number: A36214; MUID:91020984
A;Accession: A36214
A;Molecule type: DNA
A;Residues: 1-355 <FRI>
A;Cross-references: EMBL:X15983
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.8%;  Score 216;  DB 1;  Length 355;
  Best Local Similarity   58.6%;  Pred. No. 9.8e-18;
  Matches   41;  Conservative   10;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
```

```
          |  |||||:|||||||| ||:||: |:||||   : || |:    |: |: | ||: |||
Db    110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 ERIFAKAPEP 179


RESULT  12
S28360
AL1 protein - beet curly top virus
C;Species: beet curly top virus
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
C;Accession: S28360
R;Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
EMBO J. 5, 1761-1767, 1986
A;Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly to
A;Reference number: S28360
A;Accession: S28360
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-385 <STA>
A;Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.0%;  Score 213;  DB 2;  Length 385;
  Best Local Similarity   55.7%;  Pred. No. 2.4e-17;
  Matches   39;  Conservative   14;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          |: |||||:|||||||| ||:||: |:||||:| ::||||:    |: |  | ||| :|
Db    137 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 196

Qy      61 DRIFDKTPEP 70
           :|| : |:|
Db    197 QKIFQRPPDP 206


RESULT  13
JQ1870
AL1 protein - tomato mottle virus (isolate Florida)
C;Species: tomato mottle virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1870
R;Abouzid, A.M.; Polston, J.E.; Hiebert, E.
J. Gen. Virol. 73, 3225-3229, 1992
A;Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated from
A;Reference number: JQ1869; MUID:93107858
A;Accession: JQ1870
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <ABO>
A;Cross-references: GB:L14460
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             58.2%;  Score 210;  DB 1;  Length 358;
  Best Local Similarity   54.3%;  Pred. No. 5.1e-17;
  Matches   38;  Conservative   15;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
            |: ||:||:|||||||| |::||: |:|||| | : || ::    |: :: | ||: |||
Db    107 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 166

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db    167 ERIFAKAPEP 176


RESULT  14
QQCVC1
AL1 protein - tomato yellow leaf curl virus
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C;Accession: D40779
R;Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.
Virology 185, 151-161, 1991
A;Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin
A;Reference number: A40779; MUID:92024070
A;Accession: D40779
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-357 <NAV>
A;Cross-references: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             55.4%;  Score 200;  DB 1;  Length 357;
  Best Local Similarity   63.9%;  Pred. No. 7.8e-16;
  Matches   39;  Conservative    9;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLDRI 63
            :|  |:|||||||| |::||| |||||: || ||| |:    |: |: |||||:||||||
Db    111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db    171 F 171


RESULT  15
S59885
replication-associated protein C1 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
C;Accession: S59885
R;Hong, Y.; Harrison, B.D.
submitted to the EMBL Data Library, February 1995
A;Description: Nucleotide sequences from tomato leaf curl viruses from different coun
d geminiviruses.
A;Reference number: S58346
A;Accession: S59885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <HON>
A;Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA88229.1; PID:g974211
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             55.1%;  Score 199;  DB 2;  Length 360;
  Best Local Similarity   57.6%;  Pred. No. 1e-15;
  Matches   38;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLDRI 63
            :| ||:|||||||| |::||| |||:|: ||  || |:    |: :: ||||||:|||||
Db    113 FGVFQIDGRSARGGQQSANDAYAEAINSGSKAAALDILREKAPKDFVLQFHNLNANLDRI 172

Qy      64 FDKTPE 69
           |  : |
Db    173 FTPSAE 178


Search completed: October 10, 2002, 01:48:43
Job time: 561 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:47:18 ; Search time 133.15 Seconds
(without alignments)
58.394 Million cell updates/sec

Title: US-09-289-346A-4
Perfect score: 361
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	361	100.0	70	21	AAB18680	Mutant peptide der
2	346	95.8	70	21	AAB18677	Peptide fragment f
3	346	95.8	356	21	AAB18687	Amino acid sequenc
4	338	93.6	70	21	AAB18685	Mutant peptide der
5	335	92.8	70	21	AAB18688	Mutant peptide der
6	334	92.5	70	21	AAB18692	Mutant peptide der
7	332	92.0	70	21	AAB18684	Mutant peptide der
8	332	92.0	70	21	AAB18690	Mutant peptide der
9	331	91.7	70	21	AAB18678	Mutant peptide der
10	331	91.7	70	21	AAB18686	Mutant peptide der
11	330	91.4	70	21	AAB18689	Mutant peptide der
12	328	90.9	70	21	AAB18691	Mutant peptide der
13	325	90.0	70	21	AAB18681	Mutant peptide der
14	324	89.8	70	21	AAB18683	Mutant peptide der
15	322	89.2	70	21	AAB18682	Mutant peptide der
16	314	87.0	70	21	AAB18679	Mutant peptide der
17	222	61.5	359	17	AAR88870	Sardinian tomato y
18	222	61.5	359	17	AAR88871	Sardinian tomato y
19	222	61.5	359	17	AAR88872	Sardinian tomato y
20	219	60.7	353	18	AAW34338	Bean golden mosaic
21	219	60.7	353	18	AAW34332	Bean golden mosaic
22	219	60.7	353	18	AAW34333	Bean golden mosaic
23	219	60.7	353	18	AAW34334	Bean golden mosaic
24	219	60.7	353	18	AAW34335	Bean golden mosaic
25	217	60.1	353	8	AAP70407	ORF 4 gene product
26	214	59.3	361	18	AAW34336	Tomato mottle viru
27	214	59.3	361	18	AAW34324	Tomato mottle viru
28	214	59.3	361	18	AAW34325	Tomato mottle viru
29	214	59.3	361	18	AAW34326	Tomato mottle viru
30	203.5	56.4	361	8	AAP70562	Product of ORF 4 f
31	200	55.4	357	18	AAW34329	Tomato yellow leaf
32	200	55.4	357	18	AAW34330	Tomato yellow leaf
33	200	55.4	357	18	AAW34331	Tomato yellow leaf
34	198	54.8	362	19	AAW56495	Tobacco leaf curli
35	192	53.2	357	18	AAW34337	Tomato yellow leaf
36	65.5	18.1	512	19	AAW68473	HIV-1 strain YBF30
37	63.5	17.6	665	22	ABB70991	Drosophila melanog
38	61.5	17.0	163	21	AAG60400	Arabidopsis thalia
39	61.5	17.0	299	18	AAW55450	H. pylori ORF 02ae
40	61.5	17.0	1203	21	AAY83275	Candida albicans C
41	61.5	17.0	1203	21	AAY84814	Amino acid sequenc
42	61	16.9	483	22	AAG91343	C glutamicum prote
43	60.5	16.8	131	18	AAW34327	Tomato mottle viru
44	60.5	16.8	334	11	AAR03211	Amino acid sequenc
45	60	16.6	214	20	AAY37449	Amino acid sequenc

ALIGNMENTS

RESULT 1
AAB18680
ID AAB18680 standard; peptide; 70 AA.
XX
AC AAB18680;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 42
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 43
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 44
FT /note= "wild type residue replaced with Ala"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.

```
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 43-44; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             100.0%;  Score 361;  DB 21;  Length 70;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-39;
  Matches   70;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiiaaaipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   2
AAB18677
ID   AAB18677 standard; peptide; 70 AA.
XX
AC   AAB18677;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Peptide fragment from Rep protein of TGMV (amino acids 110-179).
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Tomato golden mosaic virus.
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
```



```
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 18; 73pp; English.
XX
CC   The present sequence is derived from a geminivirus replication (Rep)
CC   protein, which is also known as AL1. AL1 binds double-stranded DNA,
CC   catalyses cleavage and ligation of single-stranded DNA, and interacts
CC   with other viral and host proteins. Mutants of the AL1 protein are used
CC   to produce transgenic plants. The mutation in AL1 is present in a
CC   ribosome binding region, and expression of mutant AL1 protein imparts
CC   increased resistance to geminivirus infection in the plant. Mutant AL1
CC   proteins are useful for producing plants having increased resistance or
CC   reduced sensitivity to a geminivirus such as tomato golden mosaic virus,
CC   tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl
CC   virus, African cassava mosaic virus, Indian cassava mosaic virus, potato
CC   yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,
CC   squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or
CC   beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             95.8%;  Score 346;  DB 21;  Length 70;
  Best Local Similarity   95.7%;  Pred. No. 1.7e-37;
  Matches   67;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          ||||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db      1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   3
AAB18687
ID   AAB18687 standard; peptide; 356 AA.
XX
AC   AAB18687;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Amino acid sequence of a geminivirus replication protein of TGMV.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 354
FT                   /note= "unspecified amino acid"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
```

```
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 47-48; 73pp; English.
XX
CC   The present sequence represents a geminivirus replication (Rep)
CC   protein, which is also known as AL1. AL1 binds double-stranded DNA,
CC   catalyses cleavage and ligation of single-stranded DNA, and interacts
CC   with other viral and host proteins. Mutants of the AL1 protein are used
CC   to produce transgenic plants. The mutation in AL1 is present in a
CC   ribosome binding region, and expression of mutant AL1 protein imparts
CC   increased resistance to geminivirus infection in the plant. Mutant AL1
CC   proteins are useful for producing plants having increased resistance or
CC   reduced sensitivity to a geminivirus such as tomato golden mosaic virus,
CC   tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl
CC   virus, African cassava mosaic virus, Indian cassava mosaic virus, potato
CC   yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,
CC   squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or
CC   beet curly top virus.
XX
SQ   Sequence   356 AA;


  Query Match             95.8%;  Score 346;  DB 21;  Length 356;
  Best Local Similarity   95.7%;  Pred. No. 1.4e-36;
  Matches   67;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db     110 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 169

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     170 drifdktpep 179


RESULT   4
AAB18685
ID   AAB18685 standard; peptide; 70 AA.
XX
AC   AAB18685;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 10
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 46; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             93.6%;  Score 338;  DB 21;  Length 70;
  Best Local Similarity   94.3%;  Pred. No. 1.8e-36;
  Matches   66;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           ||||||||| ||||||||||||||||||||||||||||||   ||||||||||||||||
Db       1 tlvwgefqvagrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT   5
AAB18688
ID   AAB18688 standard; peptide; 70 AA.
XX
AC   AAB18688;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 19
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 20
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
```

```
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 48; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             92.8%;  Score 335;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 4.5e-36;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           ||||||||||||||||  ||||||||||||||||||||   ||||||||||||||||
Db       1 tlvwgefqvdgrsarggcaasndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT   6
AAB18692
ID   AAB18692 standard; peptide; 70 AA.
XX
AC   AAB18692;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 66
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 69
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 50; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             92.5%;  Score 334;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 6.1e-36;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||| || |
Db      61 drifdatpap 70


RESULT   7
AAB18684
ID   AAB18684 standard; peptide; 70 AA.
XX
AC   AAB18684;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 7
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 8
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
```

```
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             92.0%;  Score 332;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 1.1e-35;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          ||||||  |||||||||||||||||||||||||||||||   ||||||||||||||||
Db      1 tlvwgeaavdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   8
AAB18690
ID   AAB18690 standard; peptide; 70 AA.
XX
AC   AAB18690;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 27
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 30
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             92.0%;  Score 332;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 1.1e-35;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          |||||||||||||||||||||||| || ||||||||||   ||||||||||||||||
Db      1 tlvwgefqvdgrsarggcqtsndaaaaalaasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   9
AAB18678
ID   AAB18678 standard; peptide; 70 AA.
XX
AC   AAB18678;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 12
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 13
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 15
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
```

```
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 42-43; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             91.7%;  Score 331;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 1.5e-35;
  Matches   64;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |||||||||||  :| |||||||||||||||||||||||||   ||||||||||||||||
Db       1 tlvwgefqvdgaaaaggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  10
AAB18686
ID   AAB18686 standard; peptide; 70 AA.
XX
AC   AAB18686;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 24
FT                   /note= "wild type residue replaced with Leu"
FT   Misc-difference 25
FT                   /note= "wild type residue replaced with Leu"
FT   Misc-difference 26
FT                   /note= "wild type residue replaced with Leu"
XX
```

```
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 46; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             91.7%;  Score 331;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 1.5e-35;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||   ||||||||||||||   ||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndlllealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  11
AAB18689
ID   AAB18689 standard; peptide; 70 AA.
XX
AC   AAB18689;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 22
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 23
```

```
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 48-49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             91.4%;  Score 330;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 2e-35;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |||||||||||||||||||||  ||||||||||||||||||   ||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsaaaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  12
AAB18691
ID   AAB18691 standard; peptide; 70 AA.
XX
AC   AAB18691;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 34
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 35
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 36
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             90.9%;  Score 328;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 3.7e-35;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |||||||||||||||||||||||||||||||||||||||   |||||   ||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnassaaaalqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  13
AAB18681
ID   AAB18681 standard; peptide; 70 AA.
XX
AC   AAB18681;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
```

```
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 47
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 48
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 49
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 44; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             90.0%;  Score 325;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 9.1e-35;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||||||||||||||||   ||   ||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipaaalfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  14
AAB18683
ID   AAB18683 standard; peptide; 70 AA.
XX
AC   AAB18683;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 59
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 61
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 62
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             89.8%;  Score 324;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 1.2e-34;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||| |
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsal 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 aaifdktpep 70


RESULT  15
AAB18682
ID   AAB18682 standard; peptide; 70 AA.
XX
AC   AAB18682;
XX
```

DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 52
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 54
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 55
FT /note= "wild type residue replaced with Ala"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DR WPI; 2000-618851/59.
XX
PT Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT -
XX
PS Claim 53; Page 44-45; 73pp; English.
XX
CC The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
SQ Sequence 70 AA;

Query Match 89.2%; Score 322; DB 21; Length 70;
Best Local Similarity 91.4%; Pred. No. 2.2e-34;
Matches 64; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          ||||||||||||||||||||||||||||||||||||||||||   ||||||| |  |||||
Db      1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfafaalnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70
```

Search completed: October 10, 2002, 01:47:18

Job time: 527 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:01:59 ; Search time 118.38 Seconds
(without alignments)
102.295 Million cell updates/sec

Title: US-09-289-346A-5
Perfect score: 359
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	298	83.0	351	12	Q91R10	Q91r10 tomato seve
2	296	82.5	352	12	Q9E000	Q9e000 tomato rugo
3	287	79.9	226	12	Q9WHF6	Q9whf6 tomato mild
4	277	77.2	361	12	Q67574	Q67574 bean golden
5	275	76.6	225	12	Q9QDB1	Q9qdb1 cowpea gold
6	275	76.6	226	12	O09727	O09727 leonurus mo
7	271	75.5	185	12	Q98693	Q98693 sida golden
8	271	75.5	314	12	Q9ELT8	Q9elt8 sweet potat
9	271	75.5	364	12	Q9QS55	Q9qs55 sweet potat
10	264	73.5	149	12	P88975	P88975 macroptiliu
11	264	73.5	190	12	Q9Z089	Q9z089 tobacco lea
12	264	73.5	190	12	Q9Z084	Q9z084 tobacco lea
13	264	73.5	233	12	Q9YLA4	Q9yla4 macroptiliu
14	263	73.3	190	12	Q9W827	Q9w827 tobacco lea
15	263	73.3	208	12	Q9Z0C4	Q9z0c4 tobacco lea
16	263	73.3	359	12	Q91M88	Q91m88 tobacco lea
17	261	72.7	208	12	Q9Z0C0	Q9z0c0 tobacco lea
18	261	72.7	208	12	Q9Z0B8	Q9z0b8 tobacco lea
19	259	72.1	208	12	Q9Z0B6	Q9z0b6 tobacco lea
20	258	71.9	203	12	Q9Z083	Q9z083 tobacco lea
21	258	71.9	234	12	O39180	O39180 geminivirid
22	251	69.9	208	12	Q9Z0A0	Q9z0a0 tobacco lea
23	251	69.9	363	12	O72705	O72705 cotton leaf
24	251	69.9	363	12	O72719	O72719 cotton leaf
25	250	69.6	190	12	Q9Z086	Q9z086 tobacco lea
26	250	69.6	208	12	Q9Z0C6	Q9z0c6 tobacco lea
27	249	69.4	349	12	Q88888	Q88888 tomato pseu
28	249	69.4	363	12	O73577	O73577 cotton leaf
29	248	69.1	359	12	Q91B86	Q91b86 ageratum ye
30	248	69.1	360	12	Q9DXI0	Q9dxi0 ageratum ye
31	247	68.8	348	12	Q911W5	Q911w5 macroptiliu
32	246	68.5	190	12	Q9Z0A7	Q9z0a7 tobacco lea
33	246	68.5	362	12	Q9IN48	Q9in48 okra enatio
34	245	68.2	307	12	Q9IET7	Q9iet7 cotton leaf
35	245	68.2	361	12	O72723	O72723 cotton leaf
36	245	68.2	362	12	O56816	O56816 chayote mos
37	244	68.0	231	12	Q96620	Q96620 african tom
38	244	68.0	354	12	Q9IN42	Q9in42 south afric
39	244	68.0	359	12	Q9YZV4	Q9yzv4 tomato yell
40	244	68.0	359	12	Q9YZV2	Q9yzv2 tomato yell
41	244	68.0	359	12	Q9YUX7	Q9yux7 tomato yell
42	244	68.0	359	12	Q9YL27	Q9yl27 tomato yell
43	243	67.7	358	12	Q9JEA2	Q9jea2 cassava gem
44	243	67.7	359	12	Q88942	Q88942 tomato yell
45	239	66.6	362	12	Q9YYT7	Q9yyt7 althea rose

ALIGNMENTS

RESULT 1
Q91R10
ID Q91R10 PRELIMINARY; PRT; 351 AA.
AC Q91R10;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE REP PROTEIN.
GN AC1.
OS Tomato severe rugose virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=158463;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MINAS GERAIS;
RA Rezende W.L., Goulart L.R., Parreira K.S., Figueiredo J.E.F.;
RT "The full-length DNA-A nucleotide sequence of a novel tomato-infecting
RT begomovirus, Tomato severe rugose virus, in Brazil.";
RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY029750; AAK50357.1; -.
SQ SEQUENCE 351 AA; 40122 MW; 87F937A4F873B6CF CRC64;

Query Match 83.0%; Score 298; DB 12; Length 351;
Best Local Similarity 80.0%; Pred. No. 8.5e-26;
Matches 56; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: ||||:||||||||||:|||||||| ||: |||||||:|  ||||||||||
Db     111 TIEWGEFQIDGRSARGGCQTANDAAAEALNAPSKDVALQIIREKLPEKFLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| : |||
Db     171 DRIFARAPEP 180
```

RESULT 2
Q9E000
ID Q9E000 PRELIMINARY; PRT; 352 AA.

```
AC   Q9E000;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN.
GN   AC1.
OS   Tomato rugose mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=134599;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Fernandes J.J., Fontes E.P.B., Brommonschenkel S.H., Carvalho M.G.,
RA   Zambolim E.M., Zerbini F.M.;
RT   "Molecular Cloning and Characterization of Tomato rugose mosaic virus
RT   (TRMV), a Begomovirus Isolated from Tomatoes at Triangulo Mineiro,
RT   Minas Gerais, Brazil.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF291705; AAG15546.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   352 AA;  40012 MW;  47CD55838E24D613 CRC64;

  Query Match             82.5%;  Score 296;  DB 12;  Length 352;
  Best Local Similarity   80.0%;  Pred. No. 1.4e-25;
  Matches   56;  Conservative    6;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: |||||:|||||||||||:|||||||||| ||: ||||||||:|   ||||||||||
Db     111 TIEWGEFQIDGRSARGGCQTANDAAAEALNAPSKDIALQIIREKMPEKFLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| : |||
Db     171 DRIFARAPEP 180


RESULT   3
Q9WHF6
ID   Q9WHF6      PRELIMINARY;      PRT;   226 AA.
AC   Q9WHF6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   tomato mild mottle geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX   NCBI_TaxID=92943;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HN96-H5KW;
RA   Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA   Maxwell D.P.;
RT   "Molecular characterization and DNA-based detection methods for
RT   vegetable-infecting geminiviruses in Central America.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF131071; AAD33471.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     226    226
SQ   SEQUENCE   226 AA;  25941 MW;  2EA4116712871A23 CRC64;

  Query Match             79.9%;  Score 287;  DB 12;  Length 226;
  Best Local Similarity   77.1%;  Pred. No. 9.2e-25;
  Matches   54;  Conservative    9;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: |||||:|||||||| ||:|||||||||||||||::||:||:|   |||:|||:|||
Db     111 TIEWGEFQIDGRSARGGQQTANDAAAEALNASSKEEAMRIIKEKLPKEFLFQYHNLSSNL 170

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180


RESULT   4
Q67574
ID   Q67574      PRELIMINARY;      PRT;   361 AA.
AC   Q67574;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE  REPLICATIVE PROTEIN.
GN   AL1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,
RA   Maxwell D.P., Russell D.R.;
RT   "Cloning of the complete DNA genomes of four bean-infecting
RT   geminiviruses and determining their infectivity by electric discharge
RT   particle acceleration.";
RL   Phytopathology 81:980-985(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,
RA   Morales F.J., Maxwell D.P.;
RT   "Differentiation of bean-infecting geminiviruses by nucleic acid
RT   hybridization probes and aspects of bean golden mosaic in Brazil.";
RL   Plant Dis. 75:336-342(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RT   "Genetic diversity in geminiviruses causing bean golden mosaic
RT   disease: The nucleotide sequence of the infectious cloned DNA
RT   components of a Brazilian isolate of bean golden mosaic geminivirus.";
RL   Submitted (MAY-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; M88686; AAA46312.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   361 AA;  41041 MW;  0094C7ACAF06B788 CRC64;

  Query Match             77.2%;  Score 277;  DB 12;  Length 361;
  Best Local Similarity   77.6%;  Pred. No. 2.2e-23;
  Matches   52;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLDRI 63
           || |||||||||| ||:||||:|||||||||||:|||:||:|   |||:|||:||||||
Db     113 WGHFQVDGRSARGGQQTANDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
           | | |:|
Db     173 FTKAPDP 179


RESULT   5
Q9QDB1
ID   Q9QDB1      PRELIMINARY;      PRT;   225 AA.
AC   Q9QDB1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   cowpea golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
```

```
OX   NCBI_TaxID=69263;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CGMV-BR;
RA   Faria J.C.;
RT   "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT   Brazil.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF188708; AAF06318.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     225    225
SQ   SEQUENCE   225 AA;  25766 MW;  1089CB6BD8D15B5D CRC64;


  Query Match             76.6%;  Score 275;  DB 12;  Length 225;
  Best Local Similarity   77.6%;  Pred. No. 2.1e-23;
  Matches   52;  Conservative    7;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLDRI 63
           || ||:|||||||| || ||||:||||||||||:|||:||:|   |||:|||:|||||| 
Db     113 WGHFQIDGRSARGGQQTINDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
           | | |||
Db     173 FKKPPEP 179


RESULT   6
O09727
ID   O09727      PRELIMINARY;      PRT;   226 AA.
AC   O09727;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   REP.
OS   Leonurus mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=58177;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LEMV- BRAZIL 1;
RA   Faria J.C., Maxwell D.P.;
RT   "Variability in geminivirus associated with Phaseolus vulgaris in
RT   Brazil.";
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U92532; AAB51157.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     226    226
SQ   SEQUENCE   226 AA;  25617 MW;  73CDB6E766083FC5 CRC64;


  Query Match             76.6%;  Score 275;  DB 12;  Length 226;
  Best Local Similarity   77.1%;  Pred. No. 2.1e-23;
  Matches   54;  Conservative    4;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: |||||||||:||| || |||||||||  |  ||||:||:|   ||||||||||
Db     111 TVEWGEFQVDGRSSRGGQQTVNDAAAEALNAPDKRTALQIIKEKLPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180


RESULT   7
Q98693
ID   Q98693      PRELIMINARY;      PRT;   185 AA.
AC   Q98693;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   AC1.
OS   sida golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51034;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA;
RA   Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RT   "Genetic Diversity among geminiviruses associated with the weed
RT   species Sida spp, Macroptilium lathyroides, and Wissadula amplissima
RT   from Jamaica.";
RL   Plant Dis. 81:1251-1258(1997).
DR   EMBL; U67926; AAB97865.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     185    185
SQ   SEQUENCE   185 AA;  20975 MW;  3913850A025A5EE1 CRC64;


  Query Match             75.5%;  Score 271;  DB 12;  Length 185;
  Best Local Similarity   71.4%;  Pred. No. 4.8e-23;
  Matches   50;  Conservative   11;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: || ||:|||||||| ||:|||||||||: :||:||:|||||:|   |||:|||:||: 
Db      89 TIEWGVFQIDGRSARGGQQTANDAAAEALNSGTKEDALKIIREKLPERYLFQYHNLSSNI 148

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     149 DRIFSKPPEP 158


RESULT   8
Q9ELT8
ID   Q9ELT8      PRELIMINARY;      PRT;   314 AA.
AC   Q9ELT8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION ASSOCIATION PROTEIN.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF288227; AAG01006.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   314 AA;  35153 MW;  6B6220613046943F CRC64;


  Query Match             75.5%;  Score 271;  DB 12;  Length 314;
  Best Local Similarity   79.4%;  Pred. No. 8.8e-23;
```

```
  Matches   54;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
          |: |||||||||||||| ||:|||||||||| ||| ||||||||:|   |||||| |||
Db    110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy     61 DRIFDKTP 68
          ||||   |
Db    170 DRIFSPPP 177


RESULT   9
Q9QS55
ID   Q9QS55      PRELIMINARY;      PRT;   364 AA.
AC   Q9QS55;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION INITIATION PROTEIN AC1.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A.;
RT   "Cloning of a DNA-A-like genomic component of sweet potato leaf curl
RT   virus : nucleotide sequence and phylogenetic relationships.";
RL   Online Publication.
DR   EMBL; AF104036; AAD47173.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   364 AA;  40680 MW;  5F79752431A09D6E CRC64;


  Query Match             75.5%;  Score 271;  DB 12;  Length 364;
  Best Local Similarity   79.4%;  Pred. No. 1e-22;
  Matches   54;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
          |: |||||||||||||| ||:|||||||||| ||| ||||||||:|   |||||| |||
Db    110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy     61 DRIFDKTP 68
          ||||   |
Db    170 DRIFSPPP 177


RESULT  10
P88975
ID   P88975      PRELIMINARY;      PRT;   149 AA.
AC   P88975;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   AC1.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICAN;
RA   Roye M.E.;
RL   Thesis (1996), Biochemistry, University of the West Indies, Jamaica.
DR   EMBL; U75278; AAB36919.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     149    149
SQ   SEQUENCE   149 AA;  16785 MW;  E4CF5EED4C9CD508 CRC64;


  Query Match             73.5%;  Score 264;  DB 12;  Length 149;
  Best Local Similarity   70.0%;  Pred. No. 2.4e-22;
  Matches   49;  Conservative   11;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
          |: || ||:|||||||| |||||||||||: :|| |::|::||:|   |||:|||:|||
Db     52 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 111

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    112 DRIFMKDPEP 121


RESULT  11
Q9Z089
ID   Q9Z089      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z089;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA3-1, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001315; BAA34033.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21432 MW;  AAC093D1D1610FAD CRC64;


  Query Match             73.5%;  Score 264;  DB 12;  Length 190;
  Best Local Similarity   62.4%;  Pred. No. 3.1e-22;
  Matches   53;  Conservative    6;  Mismatches   10;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
          || || ||:|||||||||| :||| |||||||| ||| |||||:|   :||:|||||||
Db     85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  12
Q9Z084
ID   Q9Z084      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z084;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA5-2, PARTIAL
```

```
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA5;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001318; BAA34039.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  AAC1C2943E3F01AD CRC64;


  Query Match             73.5%;  Score 264;  DB 12;  Length 190;
  Best Local Similarity   62.4%;  Pred. No. 3.1e-22;
  Matches   53;  Conservative    6;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           || || ||:|||||||||| :||| ||||||||| ||| |||||:|   :||:||||||
Db      85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  13
Q9YLA4
ID   Q9YLA4      PRELIMINARY;      PRT;   233 AA.
AC   Q9YLA4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E.;
RT   "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
RT   from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E., McLaughlin W.A., Maxwell D.P.;
RT   "Molecular characterization of two distinct geminiviruses infecting M.
RT   lathyroides from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF098940; AAD17850.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     233    233
SQ   SEQUENCE   233 AA;  26356 MW;  AA490AF4D2166A02 CRC64;


  Query Match             73.5%;  Score 264;  DB 12;  Length 233;
  Best Local Similarity   70.0%;  Pred. No. 3.9e-22;
  Matches   49;  Conservative   11;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: || ||:|||||||| ||||||||||: :|| |::|::||:|   |||:|||:|||
Db     110 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 169

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     170 DRIFMKDPEP 179


RESULT  14
Q9W827
ID   Q9W827      PRELIMINARY;      PRT;   190 AA.
AC   Q9W827;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE GORA1-1, PARTIAL AND
DE   COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GORA1;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001303; BAA34010.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  93C3742A8EBDB7EB CRC64;


  Query Match             73.3%;  Score 263;  DB 12;  Length 190;
  Best Local Similarity   62.4%;  Pred. No. 4e-22;
  Matches   53;  Conservative    6;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           || || ||||||||||||| :||| ||||||:|| ||| |||||:|   :||:||||||
Db      85 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFSNSSFDQVPE 169


RESULT  15
Q9Z0C4
ID   Q9Z0C4      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0C4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE ABURA3-1, PARTIAL AND
DE   COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ABURA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001294; BAA33992.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
```

```
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23526 MW;  249CC31D8729C72D CRC64;


  Query Match             73.3%;  Score 263;  DB 12;  Length 208;
  Best Local Similarity   62.4%;  Pred. No. 4.5e-22;
  Matches   53;  Conservative    6;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           || || |||||||||||||| :||| |||||:|| ||| |||||:|   :||:|||||||
Db     102 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSNSSFDQVPE 186


Search completed: October 10, 2002, 02:01:59
Job time: 1092 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:09:51 ; Search time 34.62 Seconds
(without alignments)
78.289 Million cell updates/sec

Title: US-09-289-346A-5
Perfect score: 359
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	343	95.5	352	1	VAL1_TGMV	P03567 tomato gold
2	247	68.8	361	1	VAL1_PYMVV	P27258 potato yell
3	243	67.7	358	1	VAL1_CLVK	P14982 cassava lat
4	243	67.7	358	1	VAL1_CLVN	P14972 cassava lat
5	232	64.6	362	1	VAL1_TYLCA	P36279 tomato yell
6	226	63.0	359	1	VAL1_TYLCU	P38609 tomato yell
7	224	62.4	349	1	VAL1_PHUV	Q06923 pepper huas
8	219	61.0	359	1	VAL1_TYLCM	P27260 tomato yell
9	216	60.2	353	1	VAL1_BGMV	P05175 bean golden
10	216	60.2	361	1	VAL1_TMOV	Q06657 tomato mott
11	215	59.9	355	1	VAL1_ABMVW	P21947 abutilon mo
12	212	59.1	358	1	VAL1_BCTV	P14991 beet curly
13	203	56.5	357	1	VAL1_TYLCV	P27259 tomato yell
14	118	32.9	347	1	VAL1_SLCV	P29048 squash leaf
15	65	18.1	630	1	VE1_HPV66	Q80957 human papil
16	63.5	17.7	1713	1	LMA3_HUMAN	Q16787 homo sapien
17	62.5	17.4	428	1	GBA1_CANAL	P28868 candida alb
18	61	17.0	631	1	VE1_HPV30	Q05112 human papil
19	60.5	16.9	447	1	TBL2_HUMAN	Q9y4p3 homo sapien
20	60.5	16.9	1070	1	P11B_HUMAN	P42338 homo sapien
21	59.5	16.6	706	1	Y006_RICPR	Q9zed6 rickettsia
22	59.5	16.6	1053	1	UBA3_WHEAT	P31252 triticum ae
23	59.5	16.6	1610	1	CCAD_MESAU	Q99244 mesocricetu
24	59.5	16.6	2161	1	CCAD_HUMAN	Q01668 homo sapien
25	59.5	16.6	2203	1	CCAD_RAT	P27732 rattus norv
26	59	16.4	703	1	GYS2_RAT	P17625 rattus norv
27	59	16.4	1852	1	CCAS_CYPCA	P22316 cyprinus ca
28	58	16.2	387	1	Y4PF_RHISN	P55615 rhizobium s
29	58	16.2	1127	1	Y855_TREPA	O83827 treponema p
30	58	16.2	1608	1	HLYA_SERMA	P15320 serratia ma
31	57.5	16.0	367	1	LHX4_MOUSE	P53776 mus musculu
32	57.5	16.0	511	1	HUTH_VIBCH	Q9ksq4 vibrio chol
33	57	15.9	355	1	CRTB_RHOSH	P54905 rhodobacter
34	57	15.9	428	1	FLT1_MOUSE	O08917 mus musculu
35	57	15.9	432	1	PROA_DEIRA	Q9rtd9 deinococcus
36	57	15.9	524	1	MPPA_MOUSE	Q9dc61 mus musculu
37	57	15.9	524	1	MPPA_RAT	P20069 rattus norv
38	57	15.9	874	1	SLAP_BACLI	P49052 bacillus li
39	56.5	15.7	1070	1	P11B_RAT	Q9z110 rattus norv
40	56	15.6	129	1	RK12_PORPU	P51339 porphyra pu
41	56	15.6	329	1	TC3A_CAEEL	P34257 caenorhabdi
42	56	15.6	351	1	PSTS_RHILO	Q98f12 rhizobium l
43	56	15.6	427	1	FLT1_HUMAN	O75955 homo sapien
44	56	15.6	428	1	FLT1_RAT	Q9z1e1 rattus norv
45	55.5	15.5	266	1	ETFB_MYCLE	O33095 mycobacteri

ALIGNMENTS

RESULT 1
VAL1_TGMV
ID VAL1_TGMV STANDARD; PRT; 352 AA.
AC P03567;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein.
GN AC1.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10831;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;
RT "Complete nucleotide sequence of the infectious cloned DNA components
RT of tomato golden mosaic virus: potential coding regions and regulatory
RT sequences.";
RL EMBO J. 3:2197-2205(1984).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; K02029; -; NOT_ANNOTATED_CDS.
DR PIR; A04170; QQCVL1.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR ProDom; PD000736; Gemini_AL1; 1.
KW ATP-binding.
FT NP_BIND 223 230 ATP (POTENTIAL).
SQ SEQUENCE 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;

Query Match 95.5%; Score 343; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 1.3e-31;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
VAL1_PYMVV
ID VAL1_PYMVV STANDARD; PRT; 361 AA.

```
AC   P27258;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 protein.
OS   Potato yellow mosaic virus (isolate Venezuela).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10828;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91311403; PubMed=1856690;
RA   Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT   "The nucleotide sequence of the infectious cloned DNA components of
RT   potato yellow mosaic virus.";
RL   J. Gen. Virol. 72:1515-1520(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00940; BAA00782.1; -.
DR   PIR; JU0364; QQCVPT.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   361 AA;  40850 MW;  5627A33BF1264383 CRC64;
```

```
  Query Match             68.8%;  Score 247;  DB 1;  Length 361;
  Best Local Similarity   68.1%;  Pred. No. 1e-20;
  Matches   47;  Conservative   10;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: || ||:|||||||| || |||||||||: :|| |::||:||:|   |||:|||: ||
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 DRIFDKTPE 69
           |||| | ||
Db     170 DRIFMKAPE 178
```

```
RESULT   3
VAL1_CLVK
ID   VAL1_CLVK      STANDARD;      PRT;   358 AA.
AC   P14982;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (40.4 kDa protein).
GN   AC1.
OS   Cassava latent virus (strain West Kenyan 844).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10818;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Gay M.R.;
RT   "Nucleotide sequence of cassava latent virus DNA.";
RL   Nature 301:260-262(1983).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J02057; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40346 MW;  ED173E753EE92D69 CRC64;
```

```
  Query Match             67.7%;  Score 243;  DB 1;  Length 358;
  Best Local Similarity   64.3%;  Pred. No. 3e-20;
  Matches   45;  Conservative   11;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: ||:||:|||||||| |::||| |:||: || ||| :||| :|   : |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178
```

```
RESULT   4
VAL1_CLVN
ID   VAL1_CLVN      STANDARD;      PRT;   358 AA.
AC   P14972;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (40.4 kDa protein).
GN   AC1.
OS   Cassava latent virus (strain Nigerian).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10819;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90174930; PubMed=2308831;
RA   Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT   "Nucleotide sequence of the infectious cloned DNA components of
RT   African cassava mosaic virus (Nigerian strain).";
RL   Nucleic Acids Res. 18:197-198(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X17095; CAA34953.1; -.
DR   PIR; S07594; S07594.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40435 MW;  1DB16BB0CB2D5E2C CRC64;
```

```
  Query Match             67.7%;  Score 243;  DB 1;  Length 358;
  Best Local Similarity   64.3%;  Pred. No. 3e-20;
  Matches   45;  Conservative   11;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: ||:||:|||||||| |::||| |:||: || ||| :||| :|   : |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168
```

```
Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   5
VAL1_TYLCA
ID   VAL1_TYLCA     STANDARD;      PRT;   362 AA.
AC   P36279;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=36447;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93139778; PubMed=8423446;
RA   Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;
RT   "Nucleotide sequence and genome organization of tomato leaf curl
RT   geminivirus.";
RL   J. Gen. Virol. 74:147-151(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
DR   PIR; JQ1887; JQ1887.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   362 AA;  41197 MW;  343E7184B4704098 CRC64;


  Query Match             64.6%;  Score 232;  DB 1;  Length 362;
  Best Local Similarity   55.3%;  Pred. No. 5.3e-19;
  Matches   47;  Conservative    8;  Mismatches   14;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           || |||||:||||||||| |::||| |:||| || ||| ::||  |   : |||||||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   6
VAL1_TYLCU
ID   VAL1_TYLCU     STANDARD;      PRT;   359 AA.
AC   P38609;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=37139;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94256836; PubMed=8198442;
RA   Noris E., Hidalgo E., Accotto G., Moriones E.;
RT   "High similarity among the tomato yellow leaf curl virus isolates
RT   from the west Mediterranean basin: the nucleotide sequence of an
RT   infectious clone from Spain.";
RL   Arch. Virol. 135:165-170(1994).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z25751; CAA81026.1; -.
DR   PIR; S39211; S39211.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  41065 MW;  2D170A51EF80A3EC CRC64;


  Query Match             63.0%;  Score 226;  DB 1;  Length 359;
  Best Local Similarity   60.9%;  Pred. No. 2.5e-18;
  Matches   42;  Conservative    9;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || ||| :|:|  |   :  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   7
VAL1_PHUV
ID   VAL1_PHUV      STANDARD;      PRT;   349 AA.
AC   Q06923;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AL1.
OS   Pepper huasteco virus (PHV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=28349;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94015007; PubMed=8409944;
RA   Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA   Rivera-Bustamante R.F.;
RT   "Complete nucleotide sequence of pepper huasteco virus: analysis and
RT   comparison with bipartite geminiviruses.";
RL   J. Gen. Virol. 74:2225-2231(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70418; CAA49856.1; -.
DR   PIR; S31875; S31875.
DR   PIR; JQ2300; JQ2300.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (BY SIMILARITY).
SQ   SEQUENCE   349 AA;  39722 MW;  D5F4E76CD56370F4 CRC64;


  Query Match             62.4%;  Score 224;  DB 1;  Length 349;
```

```
  Best Local Similarity   61.4%;  Pred. No. 4.1e-18;
  Matches   43;  Conservative   11;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: |||||:|||||||| |::|| |:|||::| |||||||:|: |     ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   8
VAL1_TYLCM
ID   VAL1_TYLCM     STANDARD;      PRT;   359 AA.
AC   P27260;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10833;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92107660; PubMed=1840676;
RA   Kheyr-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,
RA   Gronenborn B.;
RT   "Tomato yellow leaf curl virus from Sardinia is a
RT   whitefly-transmitted monopartite geminivirus.";
RL   Nucleic Acids Res. 19:6763-6769(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X61153; CAA43466.1; -.
DR   PIR; S22593; S22593.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  40733 MW;  9717B4A07C93EFA7 CRC64;


  Query Match             61.0%;  Score 219;  DB 1;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 1.6e-17;
  Matches   40;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |   :  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   9
VAL1_BGMV
ID   VAL1_BGMV      STANDARD;      PRT;   353 AA.
AC   P05175;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   AL1 protein (40.2 kDa protein).
GN   AC1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Howarth A.J., Caton J., Bossert M., Goodman R.M.;
RT   "Nucleotide sequence of bean golden mosaic virus and a model for gene
RT   regulation in geminiviruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M10070; AAA46318.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   353 AA;  40190 MW;  80FA779DF6029A34 CRC64;


  Query Match             60.2%;  Score 216;  DB 1;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 3.3e-17;
  Matches   42;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: ||:|||||||||||| |::||: |:|||| | | || |::|: |   : | ||: |||
Db     110 TIEWGQFQVDGRSARGGQQSANDSYAKALNADSIESALTILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :|||  | |||
Db     170 ERIFFKVPEP 179


RESULT  10
VAL1_TMOV
ID   VAL1_TMOV      STANDARD;      PRT;   361 AA.
AC   Q06657;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AL1.
OS   Tomato mottle virus (isolate Florida) (TMoV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=36449;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93107858; PubMed=1469361;
RA   Abouzid A.M., Polston J.E., Hiebert E.;
RT   "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT   isolated from tomatoes in Florida.";
RL   J. Gen. Virol. 73:3225-3229(1992).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; L14460; AAC32414.1; -.
DR   PIR; JQ1870; JQ1870.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (BY SIMILARITY).
SQ   SEQUENCE   361 AA;  40516 MW;  8138B65CEEAC6950 CRC64;


  Query Match             60.2%;  Score 216;  DB 1;  Length 361;
  Best Local Similarity   57.1%;  Pred. No. 3.4e-17;
  Matches   40;  Conservative   14;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: ||:||:|||||||| |::||: |:|||| | : || ::||: |   : | ||: |||
Db     110 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
VAL1_ABMVW
ID   VAL1_ABMVW     STANDARD;      PRT;   355 AA.
AC   P21947;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AC1.
OS   Abutilon mosaic virus (isolate West India).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10816;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91020984; PubMed=2219703;
RA   Frischmuth T., Zimmat G., Jeske H.;
RT   "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT   as well as eukaryotic features.";
RL   Virology 178:461-468(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR   PIR; A36214; QQCVW1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   355 AA;  40257 MW;  16A2CA8A63251E95 CRC64;


  Query Match             59.9%;  Score 215;  DB 1;  Length 355;
  Best Local Similarity   58.6%;  Pred. No. 4.4e-17;
  Matches   41;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |  |||||:|||||||| ||:||: |:|||||   : || |::|: |   : | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :|||| | |||
Db     170 ERIFAKAPEP 179


RESULT  12
VAL1_BCTV
ID   VAL1_BCTV      STANDARD;      PRT;   358 AA.
AC   P14991;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 protein (40.8 kDa protein).
OS   Beet curly top virus (BCTV).
OC   Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
OX   NCBI_TaxID=10840;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT   "The nucleotide sequence of an infectious clone of the geminivirus
RT   beet curly top virus.";
RL   EMBO J. 5:1761-1767(1986).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40889 MW;  39A45FE3C0B9C333 CRC64;


  Query Match             59.1%;  Score 212;  DB 1;  Length 358;
  Best Local Similarity   55.7%;  Pred. No. 9.6e-17;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: |||||:|||||||| ||:||: |:||||:| ::||||::|: |     | ||| :|
Db     110 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 169

Qy      61 DRIFDKTPEP 70
           :|| : |:|
Db     170 QKIFQRPPDP 179


RESULT  13
VAL1_TYLCV
ID   VAL1_TYLCV     STANDARD;      PRT;   357 AA.
AC   P27259;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10832;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92024070; PubMed=1926771;
RA   Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;
RT   "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RT   with a single genomic component.";
RL   Virology 185:151-161(1991).
```

```
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15656; CAA33688.1; -.
DR   PIR; D40779; QQCVC1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     219    226       ATP (POTENTIAL).
SQ   SEQUENCE   357 AA;  40678 MW;  939AB68E1AB3B2A7 CRC64;


  Query Match             56.5%;  Score 203;  DB 1;  Length 357;
  Best Local Similarity   65.6%;  Pred. No. 1e-15;
  Matches   40;  Conservative    9;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLDRI 63
           :|  |:||||||||  |::||| |||||: || ||| |::|| |   : ||||||||||
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171


RESULT  14
VAL1_SLCV
ID   VAL1_SLCV      STANDARD;      PRT;   347 AA.
AC   P29048;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   AL1 protein.
OS   Squash leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10829;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91082449; PubMed=1984668;
RA   Lazarowitz S.G., Lazdins I.B.;
RT   "Infectivity and complete nucleotide sequence of the cloned genomic
RT   components of a bipartite squash leaf curl geminivirus with a broad
RT   host range phenotype.";
RL   Virology 180:58-69(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M38183; AAC32410.1; ALT_INIT.
DR   PIR; C36785; QQCVS1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     218    225       ATP (POTENTIAL).
SQ   SEQUENCE   347 AA;  39110 MW;  AFDAEBDDE122110E CRC64;


  Query Match             32.9%;  Score 118;  DB 1;  Length 347;
  Best Local Similarity   37.9%;  Pred. No. 4.2e-06;
  Matches   25;  Conservative   12;  Mismatches   25;  Indels    4;  Gaps    1;

Qy       5 GEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLDRIF 64
           |:::| |    |     :|    |:|| |  ||| ||:   |   :  :||| :|::|:|
Db     116 GQYKVSG----GSKSNKDDVYHNAVNAGSAGEALDIIKAGDPKTFIVNYHNLLANVERLF 171

Qy      65 DKTPEP 70
             | |||
Db     172 QKPPEP 177


RESULT  15
VE1_HPV66
ID   VE1_HPV66      STANDARD;      PRT;   630 AA.
AC   Q80957;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Replication protein E1.
GN   E1.
OS   Human papillomavirus type 66.
OC   Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC   Papillomavirus.
OX   NCBI_TaxID=37119;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Delius H.;
RL   Submitted (OCT-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: ATP-DEPENDENT DNA HELICASE REQUIRED FOR INITIATION OF
CC       VIRAL DNA REPLICATION. IT FORMS A COMPLEX WITH THE VIRAL E2
CC       PROTEIN. THE E1-E2 COMPLEX BINDS TO THE REPLICATION ORIGIN WHICH
CC       CONTAINS BINDING SITES FOR BOTH PROTEINS.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U31794; AAA79501.1; -.
DR   InterPro; IPR001177; Papillom_E1.
DR   Pfam; PF00519; E1; 1.
DR   Pfam; PF00524; E1_N; 1.
KW   Early protein; DNA replication; Helicase; ATP-binding; DNA-binding;
KW   Nuclear protein.
FT   NP_BIND     458    465       ATP (POTENTIAL).
SQ   SEQUENCE   630 AA;  72065 MW;  22DDDA5934F7291B CRC64;


  Query Match             18.1%;  Score 65;  DB 1;  Length 630;
  Best Local Similarity   25.5%;  Pred. No. 8.3;
  Matches   13;  Conservative   14;  Mismatches   24;  Indels    0;  Gaps    0;

Qy      19 QTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLDRIFDKTPE 69
           :|   : :  ||   ::  :|  : : || ||: :    ||:  :: :|||
Db     269 KTITKSLSSILNVPQEQMLIQPPKLRSPAVALYFYKTAMSNISEVYGETPE 319


Search completed: October 10, 2002, 02:09:52
Job time: 1350 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:48:43 ; Search time 69.63 Seconds
(without alignments)
96.600 Million cell updates/sec

Title: US-09-289-346A-5
Perfect score: 359
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	343	95.5	352	1	QQCVL1	AL1 protein - toma
2	247	68.8	361	1	QQCVPT	AL1 protein - pota
3	243	67.7	358	2	S07594	hypothetical prote
4	232	64.6	362	1	JQ1887	AL1 protein - toma
5	226	63.0	359	2	S39211	gene C1 protein -
6	224	62.4	349	2	JQ2300	replicase - pepper
7	224	62.4	349	2	S31875	AL1 protein - pepp
8	219	61.0	359	2	S22593	hypothetical prote
9	216	60.2	351	2	JQ2327	AL1 protein - Indi
10	216	60.2	358	1	JQ1870	AL1 protein - toma
11	215	59.9	355	1	QQCVW1	AV1 protein - abut
12	213	59.3	359	2	S39235	gene C1 protein -
13	212	59.1	385	2	S28360	AL1 protein - beet
14	209	58.2	360	2	S59885	replication-associ
15	203	56.5	357	1	QQCVC1	AL1 protein - toma
16	134	37.3	131	2	S45059	AC1 protein (clone
17	118	32.9	347	1	QQCVS1	AL1 protein - squa
18	65	18.1	587	2	JC1419	Fc gamma (IgG) rec
19	63.5	17.7	1033	2	E97700	hypothetical prote
20	63.5	17.7	1713	2	A55347	adhesive ligand ep
21	62.5	17.4	642	2	D90558	topoisomerase iv s
22	62	17.3	160	2	G82060	hypothetical prote
23	62	17.3	308	2	E86840	protein maturation
24	61.5	17.1	300	2	G89582	protein K07E3.3 [i
25	61.5	17.1	538	2	AD0285	probable periplasm
26	61	17.0	631	2	S36505	E1 protein - human
27	61	17.0	840	2	T36175	probable large ATP
28	60.5	16.9	338	2	AG3582	iron(III)-binding
29	60.5	16.9	429	2	A44384	GTP-binding regula
30	60.5	16.9	435	2	E59096	hypothetical prote
31	60.5	16.9	447	2	T12544	hypothetical prote
32	60.5	16.9	1070	1	A54600	1-phosphatidylinos
33	60	16.7	397	2	B71078	probable NADH oxid
34	60	16.7	714	2	C95382	probable ferrichro
35	59.5	16.6	706	2	H71707	hypothetical prote
36	59.5	16.6	1053	2	T06483	probable ubiquitin
37	59.5	16.6	1610	2	A46227	voltage-dependent
38	59.5	16.6	1646	2	JH0422	voltage-dependent
39	59.5	16.6	2137	2	T05244	hypothetical prote
40	59.5	16.6	2161	2	JH0564	calcium channel al
41	59.5	16.6	2181	2	A38198	calcium channel al
42	59.5	16.6	2203	2	T42742	voltage-dependent
43	59	16.4	351	2	E97408	fbpA (AE006039) [i
44	59	16.4	351	2	AF2626	hypothetical prote
45	59	16.4	451	2	S26839	retrovirus-related

ALIGNMENTS

RESULT 1
QQCVL1
AL1 protein - tomato golden mosaic virus
C;Species: tomato golden mosaic virus
A;Note: host Nicotiana sp. (tobacco)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C;Accession: A04170
R;Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.
EMBO J. 3, 2197-2205, 1984
A;Title: Complete nucleotide sequence of the infectious cloned DNA components of toma
A;Reference number: A04163
A;Accession: A04170
A;Molecule type: DNA
A;Residues: 1-352 <HAM>
C;Comment: The genome consists of two circular, single-stranded DNA components, DNA A
.
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 95.5%; Score 343; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 1.6e-31;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
QQCVPT
AL1 protein - potato yellow mosaic virus (isolate Venezuela)
C;Species: potato yellow mosaic virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C;Accession: JU0364
R;Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.
J. Gen. Virol. 72, 1515-1520, 1991
A;Title: The nucleotide sequence of the infectious cloned DNA components of potato ye
A;Reference number: JU0362; MUID:91311403
A;Accession: JU0364
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-361 <COU>
A;Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             68.8%;  Score 247;  DB 1;  Length 361;
  Best Local Similarity   68.1%;  Pred. No. 1.6e-20;
  Matches   47;  Conservative   10;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: || ||:|||||||| || |||||||||: :|| |::||:||:|   |||:|||: ||
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 DRIFDKTPE 69
           |||| | ||
Db     170 DRIFMKAPE 178


RESULT   3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C;Species: cassava latent virus
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C;Accession: S07594
R;Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A;Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A;Reference number: S07590; MUID:90174930
A;Accession: S07594
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <MOR>
A;Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C;Genetics:
A;Map position: segment DNA1
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             67.7%;  Score 243;  DB 2;  Length 358;
  Best Local Similarity   64.3%;  Pred. No. 4.5e-20;
  Matches   45;  Conservative   11;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: ||:||:|||||||| |::||| |:||: || ||| :||| :|   : |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   4
JQ1887
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1887
R;Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A;Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A;Reference number: JQ1885; MUID:93139778
A;Accession: JQ1887
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-362 <DRY>
A;Cross-references: GB:S53251
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             64.6%;  Score 232;  DB 1;  Length 362;
  Best Local Similarity   55.3%;  Pred. No. 8.2e-19;
  Matches   47;  Conservative    8;  Mismatches   14;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           || |||||:|||||||| |::||| |:|||  || ||| ::||  |   : |||||||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   5
S39211
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C;Accession: S39211
R;Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A;Description: High similarity among the tomato yellow leaf curl virus isolates from
A;Reference number: S39209
A;Accession: S39211
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <NOR>
A;Cross-references: EMBL:Z25751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             63.0%;  Score 226;  DB 2;  Length 359;
  Best Local Similarity   60.9%;  Pred. No. 3.9e-18;
  Matches   42;  Conservative    9;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || ||| :|:|  |   :  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   6
JQ2300
replicase - pepper huasteco virus (component A)
N;Alternate names: ORF AL1 protein
C;Species: pepper huasteco virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C;Accession: JQ2300
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A;Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A;Reference number: JQ2299; MUID:94015007
A;Accession: JQ2300
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             62.4%;  Score 224;  DB 2;  Length 349;
  Best Local Similarity   61.4%;  Pred. No. 6.5e-18;
  Matches   43;  Conservative   11;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: |||||:|||||||| |::||  |:|||::| |||||||:|: |     ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   7
S31875
AL1 protein - pepper rizado amarillo virus
C;Species: pepper rizado amarillo virus
```

```
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Sep-1999
C;Accession: S31875
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera-Bu
submitted to the EMBL Data Library, February 1993
A;Description: Complete nucleotide sequence of pepper huasteco virus: analysis and compa
A;Reference number: S31872
A;Accession: S31875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
A;Note: the source is designated as pepper huasteco virus
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             62.4%;  Score 224;  DB 2;  Length 349;
  Best Local Similarity   61.4%;  Pred. No. 6.5e-18;
  Matches   43;  Conservative   11;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: |||||:|||||||| |::||  |:|||::| |||||||:|: |     ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   8
S22593
hypothetical protein C4 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Sep-1999
C;Accession: S22593
R;Kheyr-Pour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.
Nucleic Acids Res. 19, 6763-6769, 1991
A;Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monoparti
A;Reference number: S22588; MUID:92107660
A;Accession: S22593
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-359 <KHE>
A;Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             61.0%;  Score 219;  DB 2;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 2.5e-17;
  Matches   40;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |   :  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   9
JQ2327
AL1 protein - Indian cassava mosaic virus
N;Alternate names: replication-associated protein
C;Species: Indian cassava mosaic virus
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C;Accession: JQ2327; S35883
R;Hong, Y.G.; Robinson, D.J.; Harrison, B.D.
J. Gen. Virol. 74, 2437-2443, 1993
A;Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-tran
A;Reference number: JQ2326; MUID:94065670
A;Accession: JQ2327
A;Molecule type: DNA
A;Residues: 1-351 <HON>
A;Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.2%;  Score 216;  DB 2;  Length 351;
  Best Local Similarity   62.7%;  Pred. No. 5.4e-17;
  Matches   42;  Conservative    9;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLDRI 63
           || ||:|||||||| |::||| | |||: || |||:|:||  |   |  ||:::||||||
Db     113 WGTFQIDGRSARGGQQSANDAYAAALNSGSKSEALKILRELAPRDYLRDFHHISSNLDRI 172

Qy      64 FDKTPEP 70
           | | | |
Db     173 FTKPPPP 179


RESULT  10
JQ1870
AL1 protein - tomato mottle virus (isolate Florida)
C;Species: tomato mottle virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1870
R;Abouzid, A.M.; Polston, J.E.; Hiebert, E.
J. Gen. Virol. 73, 3225-3229, 1992
A;Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated f
A;Reference number: JQ1869; MUID:93107858
A;Accession: JQ1870
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <ABO>
A;Cross-references: GB:L14460
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.2%;  Score 216;  DB 1;  Length 358;
  Best Local Similarity   57.1%;  Pred. No. 5.5e-17;
  Matches   40;  Conservative   14;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: ||:||:|||||||| |::||: |:|||| | : || ::||: |   : | ||: |||
Db     107 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 166

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     167 ERIFAKAPEP 176


RESULT  11
QQCVW1
AV1 protein - abutilon mosaic virus (isolate West India)
C;Species: abutilon mosaic virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Apr-1994
C;Accession: A36214
R;Frischmuth, T.; Zimmat, G.; Jeske, H.
Virology 178, 461-468, 1990
A;Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as
A;Reference number: A36214; MUID:91020984
A;Accession: A36214
A;Molecule type: DNA
A;Residues: 1-355 <FRI>
A;Cross-references: EMBL:X15983
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.9%;  Score 215;  DB 1;  Length 355;
  Best Local Similarity   58.6%;  Pred. No. 7.1e-17;
  Matches   41;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;
```

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |  |||||:||||||||| ||:||: |:|||||   : || |::|: |   : | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  12
S39235
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Sep-1997
C;Accession: S39235
R;Crespi, S.; Noris, E.; Vaira, A.; Bosco, D.; Accotto, G.
submitted to the EMBL Data Library, December 1993
A;Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.
A;Reference number: S39233
A;Accession: S39235
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <CRE>
A;Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.3%;  Score 213;  DB 2;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 1.2e-16;
  Matches   39;  Conservative   12;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || ||| :|::  |   :  |||::||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINARSKSEALDVIKQLAPRDYVLHFHNISSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT  13
S28360
AL1 protein - beet curly top virus
C;Species: beet curly top virus
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
C;Accession: S28360
R;Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
EMBO J. 5, 1761-1767, 1986
A;Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly to
A;Reference number: S28360
A;Accession: S28360
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-385 <STA>
A;Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.1%;  Score 212;  DB 2;  Length 385;
  Best Local Similarity   55.7%;  Pred. No. 1.7e-16;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: |||||:|||||||| ||:||: |:|||||:| ::|||||::|: |     | ||| :|
Db     137 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 196

Qy      61 DRIFDKTPEP 70
             :|| : |:|
Db     197 QKIFQRPPDP 206


RESULT  14
S59885
replication-associated protein C1 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
C;Accession: S59885
R;Hong, Y.; Harrison, B.D.
submitted to the EMBL Data Library, February 1995
A;Description: Nucleotide sequences from tomato leaf curl viruses from different coun
d geminiviruses.
A;Reference number: S58346
A;Accession: S59885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <HON>
A;Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA88229.1; PID:g974211
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             58.2%;  Score 209;  DB 2;  Length 360;
  Best Local Similarity   62.1%;  Pred. No. 3.5e-16;
  Matches   41;  Conservative   10;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLDRI 63
           :| ||:|||||||| |::||| |||:|: ||  || |:||| |   : ||||||:|||||
Db     113 FGVFQIDGRSARGGQQSANDAYAEAINSGSKAAALDILREKAPKDFVLQFHNLNANLDRI 172

Qy      64 FDKTPE 69
           |  : |
Db     173 FTPSAE 178


RESULT  15
QQCVC1
AL1 protein - tomato yellow leaf curl virus
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C;Accession: D40779
R;Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.
Virology 185, 151-161, 1991
A;Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin
A;Reference number: A40779; MUID:92024070
A;Accession: D40779
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-357 <NAV>
A;Cross-references: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             56.5%;  Score 203;  DB 1;  Length 357;
  Best Local Similarity   65.6%;  Pred. No. 1.7e-15;
  Matches   40;  Conservative    9;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLDRI 63
           :|  |:|||||||| |::||| |||||: || ||| |::|| |   : |||||:||||||
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171


Search completed: October 10, 2002, 01:48:44
Job time: 562 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:09:02 ; Search time 407.82 Seconds
(without alignments)
47.487 Million cell updates/sec

Title: US-09-289-346A-5
Perfect score: 359
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1006125 seqs, 276659714 residues

Total number of hits satisfying chosen parameters: 1006125

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	71	19.8	340	6	US-10-219-999-59588	Sequence 59588, A
2	66.5	18.5	512	5	US-09-791-537-52432	Sequence 52432, A
3	66.5	18.5	1449	5	US-09-791-537-52442	Sequence 52442, A
4	65	18.1	587	5	US-09-791-537-33198	Sequence 33198, A
5	63.5	17.7	1238	5	US-09-791-537-144585	Sequence 144585,
6	63.5	17.7	1331	5	US-09-791-537-118290	Sequence 118290,
7	63.5	17.7	1713	5	US-09-643-992A-2	Sequence 2, Appli
8	63.5	17.7	1713	5	US-09-791-537-87556	Sequence 87556, A
9	63.5	17.7	1713	6	US-10-171-311-113	Sequence 113, App
10	63.5	17.7	1806	5	US-09-791-537-53064	Sequence 53064, A
11	63.5	17.7	3298	7	US-60-365-384-181	Sequence 181, App
12	63.5	17.7	3332	7	US-60-365-384-182	Sequence 182, App
13	62.5	17.4	180	6	US-10-155-881-31455	Sequence 31455, A
14	62.5	17.4	199	6	US-10-155-881-18282	Sequence 18282, A
15	62.5	17.4	429	5	US-09-791-537-30006	Sequence 30006, A
16	62.5	17.4	434	6	US-10-179-131-5314	Sequence 5314, Ap
17	62.5	17.4	945	5	US-09-791-537-132140	Sequence 132140,
18	61.5	17.1	300	5	US-09-791-537-98310	Sequence 98310, A
19	61.5	17.1	300	7	US-60-360-039-6839	Sequence 6839, Ap
20	61.5	17.1	365	6	US-10-155-881-28183	Sequence 28183, A
21	61.5	17.1	365	7	US-60-411-837-483	Sequence 483, App
22	61.5	17.1	475	7	US-60-389-987-1438	Sequence 1438, Ap
23	61.5	17.1	476	5	US-09-582-687A-62	Sequence 62, Appl
24	60.5	16.9	429	5	US-09-791-537-53935	Sequence 53935, A
25	60.5	16.9	435	5	US-09-791-537-147074	Sequence 147074,
26	60.5	16.9	447	5	US-09-991-150-369	Sequence 369, App
27	60.5	16.9	447	5	US-09-989-328-369	Sequence 369, App
28	60.5	16.9	447	6	US-10-103-295-129	Sequence 129, App
29	60.5	16.9	447	6	US-10-121-062-254	Sequence 254, App
30	60.5	16.9	447	6	US-10-173-689-254	Sequence 254, App
31	60.5	16.9	447	6	US-10-173-690-254	Sequence 254, App
32	60.5	16.9	447	6	US-10-173-691-254	Sequence 254, App
33	60.5	16.9	447	6	US-10-173-692-254	Sequence 254, App
34	60.5	16.9	447	6	US-10-173-695-254	Sequence 254, App
35	60.5	16.9	447	6	US-10-173-697-254	Sequence 254, App
36	60.5	16.9	447	6	US-10-173-698-254	Sequence 254, App
37	60.5	16.9	447	6	US-10-173-699-254	Sequence 254, App
38	60.5	16.9	447	6	US-10-173-701-254	Sequence 254, App
39	60.5	16.9	447	6	US-10-173-704-254	Sequence 254, App
40	60.5	16.9	447	6	US-10-173-708-254	Sequence 254, App
41	60.5	16.9	447	6	US-10-174-571-254	Sequence 254, App
42	60.5	16.9	447	6	US-10-174-574-254	Sequence 254, App
43	60.5	16.9	447	6	US-10-174-583-254	Sequence 254, App
44	60.5	16.9	447	6	US-10-173-694-254	Sequence 254, App
45	60.5	16.9	447	6	US-10-173-707-254	Sequence 254, App

ALIGNMENTS

RESULT 1
US-10-219-999-59588
; Sequence 59588, Application US/10219999
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: cDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)C
; CURRENT APPLICATION NUMBER: US/10/219,999
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/324,109
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/312,544
; PRIOR FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 63520
; SEQ ID NO 59588
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Zea mays
US-10-219-999-59588

Query Match 19.8%; Score 71; DB 6; Length 340;
Best Local Similarity 28.2%; Pred. No. 1.9;
Matches 20; Conservative 13; Mismatches 22; Indels 16; Gaps 3;

```
Qy      15 RGGCQTSNDAAAEAL-NASSKEEALQ--------IIREKIPAAALFQFHN-------LNS 58
           | || :    |  |  : ::|::  |        :::|::| |  |  ||       |||
Db      30 RPGCSSRTPTATSATHHGATKQQVRQGKVSDETKMVQEELPQAWWFDSHNLARPSPWLNS 89

Qy      59 NLDRIFDKTPE 69
             |  : ||| :
Db      90 TLSELDDKTKQ 100
```

RESULT 2
US-09-791-537-52432
; Sequence 52432, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210

```
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 52432
;   LENGTH: 512
;   TYPE: PRT
;   ORGANISM: Human immunodeficiency virus type 1
US-09-791-537-52432


  Query Match             18.5%;  Score 66.5;  DB 5;  Length 512;
  Best Local Similarity   28.6%;  Pred. No. 12;
  Matches   24;  Conservative   13;  Mismatches   30;  Indels   17;  Gaps    4;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPA-----AALF----- 51
           |||   :::  :   |   : |  |: ||    | ||:   |:: :     | |:
Db      34 LVWASRELERFACNPGLMDTADGCAKLLN--QLEPALKTGSEELRSLYNALAVLYCVHSR 91
.
Qy      52 -QFHNLNSNLDRIFDK----TPEP 70
            | ||    ||:| :|     |||
Db      92 IQIHNTQEALDKIKEKQEQHKPEP 115


RESULT   3
US-09-791-537-52442
; Sequence 52442, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 52442
;   LENGTH: 1449
;   TYPE: PRT
;   ORGANISM: Human immunodeficiency virus type 1
US-09-791-537-52442


  Query Match             18.5%;  Score 66.5;  DB 5;  Length 1449;
  Best Local Similarity   28.6%;  Pred. No. 44;
  Matches   24;  Conservative   13;  Mismatches   30;  Indels   17;  Gaps    4;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPA-----AALF----- 51
           |||   :::  :   |   : |  |: ||    | ||:   |:: :     | |:
Db      34 LVWASRELERFACNPGLMDTADGCAKLLN--QLEPALKTGSEELRSLYNALAVLYCVHSR 91

Qy      52 -QFHNLNSNLDRIFDK----TPEP 70
            | ||    ||:| :|     |||
Db      92 IQIHNTQEALDKIKEKQEQHKPEP 115


RESULT   4
US-09-791-537-33198
; Sequence 33198, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 33198
;   LENGTH: 587
;   TYPE: PRT
;   ORGANISM: Streptococcus sp
US-09-791-537-33198


  Query Match             18.1%;  Score 65;  DB 5;  Length 587;
  Best Local Similarity   30.0%;  Pred. No. 22;
  Matches   18;  Conservative   12;  Mismatches   22;  Indels    8;  Gaps    1;

Qy      12 RSARGGCQTSNDAAAEALNASSKE--------EALQIIREKIPAAALFQFHNLNSNLDRI 63
           :|| || :|: |  |: || :|::         :|||   |     ||   :| |  : ::
Db     114 KSANGGLKTALDDMADTLNQASRKVNELSAQNKALQAEAEAAAQKALDALNNKNEQIAKL 173


RESULT   5
US-09-791-537-144585
; Sequence 144585, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 144585
;   LENGTH: 1238
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-791-537-144585


  Query Match             17.7%;  Score 63.5;  DB 5;  Length 1238;
  Best Local Similarity   34.4%;  Pred. No. 87;
  Matches   21;  Conservative   12;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPAAALFQFHNLNSNLDRIFD 65
           ||  :|      :  || :| |  ||: | ||| :|:| :|  |      |:|| |:: :
Db     553 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRKA----KTLSSNSDKLLN 608

Qy      66 K 66
           :
Db     609 E 609


RESULT   6
US-09-791-537-118290
; Sequence 118290, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 118290
;   LENGTH: 1331
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-791-537-118290


  Query Match             17.7%;  Score 63.5;  DB 5;  Length 1331;
  Best Local Similarity   34.4%;  Pred. No. 95;
  Matches   21;  Conservative   12;  Mismatches   21;  Indels    7;  Gaps    3;
```

```
Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPAAALFQFHNLNSNLDRIFD 65
           ||  :|      :  || :| |  ||: | ||| :|:| :|  |      |:|| |:: :
Db     646 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRKA----KTLSSNSDKLLN 701

Qy      66 K 66
           :
Db     702 E 702
```

RESULT 7
US-09-643-992A-2
; Sequence 2, Application US/09643992A
; GENERAL INFORMATION:
; APPLICANT: Oriental Yeast Co., Ltd.
; APPLICANT: NAITOU Toshikuni
; TITLE OF INVENTION: Method for neuranagenesis
; FILE REFERENCE: FP/O-21-8US
; CURRENT APPLICATION NUMBER: US/09/643,992A
; CURRENT FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 2
; LENGTH: 1713
; TYPE: PRT
; ORGANISM: Human
US-09-643-992A-2

```
  Query Match             17.7%;  Score 63.5;  DB 5;  Length 1713;
  Best Local Similarity   34.4%;  Pred. No. 1.3e+02;
  Matches   21;  Conservative   12;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPAAALFQFHNLNSNLDRIFD 65
           ||  :|      :  || :| |  ||: | ||| :|:| :|  |      |:|| |:: :
Db     553 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRKA----KTLSSNSDKLLN 608

Qy      66 K 66
           :
Db     609 E 609
```

RESULT 8
US-09-791-537-87556
; Sequence 87556, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 87556
; LENGTH: 1713
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-87556

```
  Query Match             17.7%;  Score 63.5;  DB 5;  Length 1713;
  Best Local Similarity   34.4%;  Pred. No. 1.3e+02;
  Matches   21;  Conservative   12;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPAAALFQFHNLNSNLDRIFD 65
           ||  :|      :  || :| |  ||: | ||| :|:| :|  |      |:|| |:: :
Db     553 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRKA----KTLSSNSDKLLN 608

Qy      66 K 66
           :
Db     609 E 609
```

RESULT 9
US-10-171-311-113
; Sequence 113, Application US/10171311
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersh, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 113
; LENGTH: 1713
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-113

```
  Query Match             17.7%;  Score 63.5;  DB 6;  Length 1713;
  Best Local Similarity   34.4%;  Pred. No. 1.3e+02;
  Matches   21;  Conservative   12;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPAAALFQFHNLNSNLDRIFD 65
           ||  :|      :  || :| |  ||: | ||| :|:| :|  |      |:|| |:: :
Db     553 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRKA----KTLSSNSDKLLN 608

Qy      66 K 66
           :
Db     609 E 609
```

RESULT 10
US-09-791-537-53064
; Sequence 53064, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 53064
; LENGTH: 1806
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-53064

```
  Query Match             17.7%;  Score 63.5;  DB 5;  Length 1806;
  Best Local Similarity   34.4%;  Pred. No. 1.4e+02;
  Matches   21;  Conservative   12;  Mismatches   21;  Indels    7;  Gaps    3;
```

```
Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPAAALFQFHNLNSNLDRIFD 65
           ||  :|      :  || :| |  ||: | ||| :|:| :|  |      |:|| |:: :
Db     646 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRKA----KTLSSNSDKLLN 701

Qy      66 K 66
           :
Db     702 E 702
```

RESULT 11
US-60-365-384-181
; Sequence 181, Application US/60365384
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Wang, Jian-rui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ghosh, Malabika
; APPLICANT: Zhao, Qing A.
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Weng, Gezhi
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 814
; CURRENT APPLICATION NUMBER: US/60/365,384
; CURRENT FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 666
; SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 181
; LENGTH: 3298
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-365-384-181

Query Match 17.7%; Score 63.5; DB 7; Length 3298;
Best Local Similarity 34.4%; Pred. No. 3e+02;
Matches 21; Conservative 12; Mismatches 21; Indels 7; Gaps 3;

```
Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPAAALFQFHNLNSNLDRIFD 65
           ||  :|      :  || :| |  ||: | ||| :|:| :|  |      |:|| |:: :
Db    2138 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRKA----KTLSSNSDKLLN 2193

Qy      66 K 66
           :
Db    2194 E 2194
```

RESULT 12
US-60-365-384-182
; Sequence 182, Application US/60365384
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Wang, Jian-rui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ghosh, Malabika
; APPLICANT: Zhao, Qing A.
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Weng, Gezhi
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 814
; CURRENT APPLICATION NUMBER: US/60/365,384
; CURRENT FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 666
; SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 182
; LENGTH: 3332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-365-384-182

Query Match 17.7%; Score 63.5; DB 7; Length 3332;
Best Local Similarity 34.4%; Pred. No. 3e+02;
Matches 21; Conservative 12; Mismatches 21; Indels 7; Gaps 3;

```
Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPAAALFQFHNLNSNLDRIFD 65
           ||  :|      :  || :| |  ||: | ||| :|:| :|  |      |:|| |:: :
Db    2172 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRKA----KTLSSNSDKLLN 2227

Qy      66 K 66
           :
Db    2228 E 2228
```

RESULT 13
US-10-155-881-31455
; Sequence 31455, Application US/10155881
; GENERAL INFORMATION:
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Lutfiyya, Linda L.
; APPLICANT: McIninch, James
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: TRANSCRIPTION IN PLANTS
; FILE REFERENCE: 38-21(15300)J
; CURRENT APPLICATION NUMBER: US/10/155,881
; CURRENT FILING DATE: 2002-05-22
; NUMBER OF SEQ ID NOS: 37595
; SEQ ID NO 31455
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Glycine max
US-10-155-881-31455

Query Match 17.4%; Score 62.5; DB 6; Length 180;
Best Local Similarity 34.3%; Pred. No. 10;
Matches 23; Conservative 9; Mismatches 24; Indels 11; Gaps 3;

```
Qy       4 WGEFQVDGR-SARGGC-------QTSNDAAAEALNASSK---EEALQIIREKIPAAALFQ 52
           ||::  : | ||| |         ||: :||     |: |    :||     :| ||
Db      80 WGKYAAEIRDSARHGARIWLGTFQTAEEAAMAYDRAAFKMRGSKALLNFPAEIVAATELS 139

Qy      53 FHNLNSN 59
           |:| |||
Db     140 FNNTNSN 146
```

RESULT 14
US-10-155-881-18282
; Sequence 18282, Application US/10155881
; GENERAL INFORMATION:
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Lutfiyya, Linda L.
; APPLICANT: McIninch, James
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH

```
;  TITLE OF INVENTION:  TRANSCRIPTION IN PLANTS
;  FILE REFERENCE: 38-21(15300)J
;  CURRENT APPLICATION NUMBER: US/10/155,881
;  CURRENT FILING DATE:  2002-05-22
;  NUMBER OF SEQ ID NOS: 37595
; SEQ ID NO 18282
;   LENGTH: 199
;   TYPE: PRT
;   ORGANISM: Glycine max
US-10-155-881-18282


  Query Match            17.4%;  Score 62.5;  DB 6;  Length 199;
  Best Local Similarity   34.3%;  Pred. No. 12;
  Matches   23;  Conservative    9;  Mismatches   24;  Indels   11;  Gaps    3;

Qy       4 WGEFQVDGR-SARGGC-------QTSNDAAAEALNASSK---EEALQIIREKIPAAALFQ 52
           ||::  : | ||| |        ||: :||     |: |    :||     :| ||
Db      99 WGKYAAEIRDSARHGARIWLGTFQTAEEAAMAYDRAAFKMRGSKALLNFPAEIVAATELS 158

Qy      53 FHNLNSN 59
           |:| |||
Db     159 FNNTNSN 165


RESULT  15
US-09-791-537-30006
; Sequence 30006, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30006
;   LENGTH: 429
;   TYPE: PRT
;   ORGANISM: Candida albicans
US-09-791-537-30006


  Query Match            17.4%;  Score 62.5;  DB 5;  Length 429;
  Best Local Similarity   29.0%;  Pred. No. 31;
  Matches   20;  Conservative   16;  Mismatches   24;  Indels    9;  Gaps    3;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           | :||:   |  |       ::||  :|| :|  ::: |  |  | | |: :   |:| :
Db     162 TDIWGK---DDDS-----NINSDAINQALESSLNKDSEQFTRLSI-AEAIHKLWKLDSGI 212

Qy      61 DRIFDKTPE 69
            : ||:: |
Db     213 KKCFDRSNE 221


Search completed: October 10, 2002, 02:09:02
Job time: 1445 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:47:18 ; Search time 133.15 Seconds
(without alignments)
58.394 Million cell updates/sec

Title: US-09-289-346A-5
Perfect score: 359
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	359	100.0	70	21	AAB18681	Mutant peptide der
2	343	95.5	70	21	AAB18677	Peptide fragment f
3	343	95.5	356	21	AAB18687	Amino acid sequenc
4	335	93.3	70	21	AAB18685	Mutant peptide der
5	332	92.5	70	21	AAB18688	Mutant peptide der
6	331	92.2	70	21	AAB18692	Mutant peptide der
7	329	91.6	70	21	AAB18684	Mutant peptide der
8	329	91.6	70	21	AAB18690	Mutant peptide der
9	328	91.4	70	21	AAB18678	Mutant peptide der
10	328	91.4	70	21	AAB18686	Mutant peptide der
11	327	91.1	70	21	AAB18689	Mutant peptide der
12	325	90.5	70	21	AAB18680	Mutant peptide der
13	325	90.5	70	21	AAB18691	Mutant peptide der
14	321	89.4	70	21	AAB18683	Mutant peptide der
15	319	88.9	70	21	AAB18682	Mutant peptide der
16	311	86.6	70	21	AAB18679	Mutant peptide der
17	220	61.3	361	18	AAW34336	Tomato mottle viru
18	220	61.3	361	18	AAW34324	Tomato mottle viru
19	220	61.3	361	18	AAW34325	Tomato mottle viru
20	220	61.3	361	18	AAW34326	Tomato mottle viru
21	219	61.0	359	17	AAR88870	Sardinian tomato y
22	219	61.0	359	17	AAR88871	Sardinian tomato y
23	219	61.0	359	17	AAR88872	Sardinian tomato y
24	218	60.7	353	18	AAW34338	Bean golden mosaic
25	218	60.7	353	18	AAW34332	Bean golden mosaic
26	218	60.7	353	18	AAW34333	Bean golden mosaic
27	218	60.7	353	18	AAW34334	Bean golden mosaic
28	218	60.7	353	18	AAW34335	Bean golden mosaic
29	216	60.2	353	8	AAP70407	ORF 4 gene product
30	208.5	58.1	361	8	AAP70562	Product of ORF 4 f
31	208	57.9	362	19	AAW56495	Tobacco leaf curli
32	203	56.5	357	18	AAW34329	Tomato yellow leaf
33	203	56.5	357	18	AAW34330	Tomato yellow leaf
34	203	56.5	357	18	AAW34331	Tomato yellow leaf
35	195	54.3	357	18	AAW34337	Tomato yellow leaf
36	66.5	18.5	512	19	AAW68473	HIV-1 strain YBF30
37	65.5	18.2	665	22	ABB70991	Drosophila melanog
38	63.5	17.7	1693	21	AAB48457	Human laminin 5 po
39	63.5	17.7	1693	21	AAB48459	Human laminin 5 po
40	63.5	17.7	1713	16	AAR70148	Deduced sequence o
41	63.5	17.7	1713	21	AAB48458	Human laminin 5 po
42	63.5	17.7	1724	21	AAB48456	Human laminin 5 po
43	63	17.5	599	22	ABG14172	Novel human diagno
44	63	17.5	599	22	AAU32811	Novel human secret
45	62.5	17.4	945	22	ABB65231	Drosophila melanog

ALIGNMENTS

RESULT 1
AAB18681
ID AAB18681 standard; peptide; 70 AA.
XX
AC AAB18681;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 47
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 48
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 49
FT /note= "wild type residue replaced with Ala"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.

```
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 44; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             100.0%;  Score 359;  DB 21;  Length 70;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-39;
  Matches   70;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipaaalfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   2
AAB18677
ID   AAB18677 standard; peptide; 70 AA.
XX
AC   AAB18677;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Peptide fragment from Rep protein of TGMV (amino acids 110-179).
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Tomato golden mosaic virus.
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 18; 73pp; English.
XX
CC   The present sequence is derived from a geminivirus replication (Rep)
CC   protein, which is also known as AL1. AL1 binds double-stranded DNA,
CC   catalyses cleavage and ligation of single-stranded DNA, and interacts
CC   with other viral and host proteins. Mutants of the AL1 protein are used
CC   to produce transgenic plants. The mutation in AL1 is present in a
CC   ribosome binding region, and expression of mutant AL1 protein imparts
CC   increased resistance to geminivirus infection in the plant. Mutant AL1
CC   proteins are useful for producing plants having increased resistance or
CC   reduced sensitivity to a geminivirus such as tomato golden mosaic virus,
CC   tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl
CC   virus, African cassava mosaic virus, Indian cassava mosaic virus, potato
CC   yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,
CC   squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or
CC   beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             95.5%;  Score 343;  DB 21;  Length 70;
  Best Local Similarity   95.7%;  Pred. No. 4.3e-37;
  Matches   67;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||
Db      1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   3
AAB18687
ID   AAB18687 standard; peptide; 356 AA.
XX
AC   AAB18687;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Amino acid sequence of a geminivirus replication protein of TGMV.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 354
FT                   /note= "unspecified amino acid"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
```

```
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 47-48; 73pp; English.
XX
CC   The present sequence represents a geminivirus replication (Rep)
CC   protein, which is also known as AL1. AL1 binds double-stranded DNA,
CC   catalyses cleavage and ligation of single-stranded DNA, and interacts
CC   with other viral and host proteins. Mutants of the AL1 protein are used
CC   to produce transgenic plants. The mutation in AL1 is present in a
CC   ribosome binding region, and expression of mutant AL1 protein imparts
CC   increased resistance to geminivirus infection in the plant. Mutant AL1
CC   proteins are useful for producing plants having increased resistance or
CC   reduced sensitivity to a geminivirus such as tomato golden mosaic virus,
CC   tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl
CC   virus, African cassava mosaic virus, Indian cassava mosaic virus, potato
CC   yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,
CC   squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or
CC   beet curly top virus.
XX
SQ   Sequence   356 AA;


  Query Match             95.5%;  Score 343;  DB 21;  Length 356;
  Best Local Similarity   95.7%;  Pred. No. 3.7e-36;
  Matches   67;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||
Db     110 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 169

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     170 drifdktpep 179


RESULT   4
AAB18685
ID   AAB18685 standard; peptide; 70 AA.
XX
AC   AAB18685;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 10
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 46; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             93.3%;  Score 335;  DB 21;  Length 70;
  Best Local Similarity   94.3%;  Pred. No. 4.8e-36;
  Matches   66;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           ||||||||| ||||||||||||||||||||||||||||||||||   ||||||||||
Db       1 tlvwgefqvagrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT   5
AAB18688
ID   AAB18688 standard; peptide; 70 AA.
XX
AC   AAB18688;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 19
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 20
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
```

```
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 48; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             92.5%;  Score 332;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 1.2e-35;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           ||||||||||||||||||  |||||||||||||||||||||||||   ||||||||||
Db       1 tlvwgefqvdgrsarggcaasndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT   6
AAB18692
ID   AAB18692 standard; peptide; 70 AA.
XX
AC   AAB18692;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 66
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 69
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 50; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             92.2%;  Score 331;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 1.6e-35;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||||||||||||   ||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           |||||  || |
Db      61 drifdatpap 70


RESULT   7
AAB18684
ID   AAB18684 standard; peptide; 70 AA.
XX
AC   AAB18684;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 7
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 8
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
```

```
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             91.6%;  Score 329;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 2.9e-35;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
          ||||||  ||||||||||||||||||||||||||||||||||||||   |||||||||||
Db      1 tlvwgeaavdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   8
AAB18690
ID   AAB18690 standard; peptide; 70 AA.
XX
AC   AAB18690;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 27
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 30
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             91.6%;  Score 329;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 2.9e-35;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
          |||||||||||||||||||||||| || ||||||||||||||   ||||||||||
Db      1 tlvwgefqvdgrsarggcqtsndaaaaalaasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   9
AAB18678
ID   AAB18678 standard; peptide; 70 AA.
XX
AC   AAB18678;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 12
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 13
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 15
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
```

```
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 42-43; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             91.4%;  Score 328;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 3.9e-35;
  Matches   64;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |||||||||||  :| ||||||||||||||||||||||||||||||||   ||||||||||
Db       1 tlvwgefqvdgaaaaggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  10
AAB18686
ID   AAB18686 standard; peptide; 70 AA.
XX
AC   AAB18686;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 24
FT                   /note= "wild type residue replaced with Leu"
FT   Misc-difference 25
FT                   /note= "wild type residue replaced with Leu"
FT   Misc-difference 26
FT                   /note= "wild type residue replaced with Leu"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 46; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             91.4%;  Score 328;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 3.9e-35;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           ||||||||||||||||||||   ||||||||||||||||||   ||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndlllealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  11
AAB18689
ID   AAB18689 standard; peptide; 70 AA.
XX
AC   AAB18689;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 22
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 23
```

```
FT                    /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 48-49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             91.1%;  Score 327;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 5.3e-35;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |||||||||||||||||||  |||||||||||||||||||||   ||||||||||
Db       1 tlvwgefqvdgrsarggcqtsaaaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  12
AAB18680
ID   AAB18680 standard; peptide; 70 AA.
XX
AC   AAB18680;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 42
FT                    /note= "wild type residue replaced with Ala"
FT   Misc-difference 43
FT                    /note= "wild type residue replaced with Ala"
FT   Misc-difference 44
FT                    /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 43-44; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             90.5%;  Score 325;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 9.6e-35;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |||||||||||||||||||||||||||||||||||||||||||||||||   ||   ||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiiaaaipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  13
AAB18691
ID   AAB18691 standard; peptide; 70 AA.
XX
AC   AAB18691;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
```

```
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 34
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 35
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 36
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             90.5%;  Score 325;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 9.6e-35;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||   ||||||||||   ||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnassaaaalqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  14
AAB18683
ID   AAB18683 standard; peptide; 70 AA.
XX
AC   AAB18683;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 59
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 61
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 62
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             89.4%;  Score 321;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 3.2e-34;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||| |
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsal 60

Qy      61 DRIFDKTPEP 70
           ||||||||
Db      61 aaifdktpep 70


RESULT  15
AAB18682
ID   AAB18682 standard; peptide; 70 AA.
XX
AC   AAB18682;
XX
```

```
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 52
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 54
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 55
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 44-45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             88.9%;  Score 319;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 5.8e-34;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||   || |  |||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfafaalnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


Search completed: October 10, 2002, 01:47:18

Job time: 527 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:01:59 ; Search time 118.38 Seconds
(without alignments)
102.295 Million cell updates/sec

Title: US-09-289-346A-6
Perfect score: 357
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FAFAALNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	291	81.5	351	12	Q91R10	Q91r10 tomato seve
2	289	81.0	352	12	Q9E000	Q9e000 tomato rugo
3	272	76.2	226	12	O09727	O09727 leonurus mo
4	272	76.2	226	12	Q9WHF6	Q9whf6 tomato mild
5	270	75.6	361	12	Q67574	Q67574 bean golden
6	268	75.1	225	12	Q9QDB1	Q9qdb1 cowpea gold
7	268	75.1	314	12	Q9ELT8	Q9elt8 sweet potat
8	268	75.1	364	12	Q9QS55	Q9qs55 sweet potat
9	265	74.2	185	12	Q98693	Q98693 sida golden
10	257	72.0	149	12	P88975	P88975 macroptiliu
11	257	72.0	233	12	Q9YLA4	Q9yla4 macroptiliu
12	251	70.3	234	12	O39180	O39180 geminivirid
13	248	69.5	190	12	Q9Z089	Q9z089 tobacco lea
14	248	69.5	190	12	Q9Z084	Q9z084 tobacco lea
15	247	69.2	190	12	Q9W827	Q9w827 tobacco lea
16	247	69.2	208	12	Q9Z0C4	Q9z0c4 tobacco lea
17	247	69.2	359	12	Q91M88	Q91m88 tobacco lea
18	245	68.6	208	12	Q9Z0C0	Q9z0c0 tobacco lea
19	245	68.6	208	12	Q9Z0B8	Q9z0b8 tobacco lea
20	243	68.1	208	12	Q9Z0B6	Q9z0b6 tobacco lea
21	242	67.8	203	12	Q9Z083	Q9z083 tobacco lea
22	242	67.8	349	12	Q88888	Q88888 tomato pseu
23	242	67.8	363	12	O73577	O73577 cotton leaf
24	239	66.9	208	12	Q9Z0A0	Q9z0a0 tobacco lea
25	238	66.7	208	12	Q9Z0C6	Q9z0c6 tobacco lea
26	237	66.4	359	12	Q91B86	Q91b86 ageratum ye
27	237	66.4	360	12	Q9DXI0	Q9dxi0 ageratum ye
28	235	65.8	363	12	O72705	O72705 cotton leaf
29	235	65.8	363	12	O72719	O72719 cotton leaf
30	234	65.5	190	12	Q9Z086	Q9z086 tobacco lea
31	233	65.3	362	12	O56816	O56816 chayote mos
32	232	65.0	359	12	Q9YZV4	Q9yzv4 tomato yell
33	232	65.0	359	12	Q9YZV2	Q9yzv2 tomato yell
34	232	65.0	359	12	Q9YUX7	Q9yux7 tomato yell
35	232	65.0	359	12	Q9YL27	Q9yl27 tomato yell
36	231	64.7	348	12	Q911W5	Q911w5 macroptiliu
37	231	64.7	353	12	O72692	O72692 beet curly
38	231	64.7	354	12	Q65438	Q65438 beet curly
39	231	64.7	359	12	Q88942	Q88942 tomato yell
40	230	64.4	190	12	Q9Z0A7	Q9z0a7 tobacco lea
41	230	64.4	362	12	Q9IN48	Q9in48 okra enatio
42	229	64.1	307	12	Q9IET7	Q9iet7 cotton leaf
43	229	64.1	361	12	O72723	O72723 cotton leaf
44	228	63.9	231	12	Q96620	Q96620 african tom
45	228	63.9	354	12	Q9IN42	Q9in42 south afric

ALIGNMENTS

RESULT 1
Q91R10
ID Q91R10 PRELIMINARY; PRT; 351 AA.
AC Q91R10;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE REP PROTEIN.
GN AC1.
OS Tomato severe rugose virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=158463;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MINAS GERAIS;
RA Rezende W.L., Goulart L.R., Parreira K.S., Figueiredo J.E.F.;
RT "The full-length DNA-A nucleotide sequence of a novel tomato-infecting
RT begomovirus, Tomato severe rugose virus, in Brazil.";
RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY029750; AAK50357.1; -.
SQ SEQUENCE 351 AA; 40122 MW; 87F937A4F873B6CF CRC64;

Query Match 81.5%; Score 291; DB 12; Length 351;
Best Local Similarity 78.6%; Pred. No. 2.6e-25;
Matches 55; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: |||||:||||||||||||:|||||||||| ||: ||||||||:|||:|| |  |||||
Db     111 TIEWGEFQIDGRSARGGCQTANDAAAEALNAPSKDVALQIIREKLPEKFLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| : |||
Db     171 DRIFARAPEP 180
```

RESULT 2
Q9E000
ID Q9E000 PRELIMINARY; PRT; 352 AA.

```
AC   Q9E000;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN.
GN   AC1.
OS   Tomato rugose mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=134599;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Fernandes J.J., Fontes E.P.B., Brommonschenkel S.H., Carvalho M.G.,
RA   Zambolim E.M., Zerbini F.M.;
RT   "Molecular Cloning and Characterization of Tomato rugose mosaic virus
RT   (TRMV), a Begomovirus Isolated from Tomatoes at Triangulo Mineiro,
RT   Minas Gerais, Brazil.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF291705; AAG15546.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   352 AA;  40012 MW;  47CD55838E24D613 CRC64;


  Query Match             81.0%;  Score 289;  DB 12;  Length 352;
  Best Local Similarity   78.6%;  Pred. No. 4.3e-25;
  Matches   55;  Conservative    7;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: |||||:|||||||||||:|||||||||| ||: |||||||:|||:|| |  |||||
Db     111 TIEWGEFQIDGRSARGGCQTANDAAAEALNAPSKDIALQIIREKMPEKFLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| : |||
Db     171 DRIFARAPEP 180


RESULT   3
O09727
ID   O09727      PRELIMINARY;      PRT;   226 AA.
AC   O09727;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   REP.
OS   Leonurus mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=58177;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LEMV- BRAZIL 1;
RA   Faria J.C., Maxwell D.P.;
RT   "Variability in geminivirus associated with Phaseolus vulgaris in
RT   Brazil.";
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U92532; AAB51157.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     226    226
SQ   SEQUENCE   226 AA;  25617 MW;  73CDB6E766083FC5 CRC64;


  Query Match             76.2%;  Score 272;  DB 12;  Length 226;
  Best Local Similarity   77.1%;  Pred. No. 2.3e-23;
  Matches   54;  Conservative    4;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: |||||||||:||| || |||||||||| |  |||||:||:|||||| |  |||||
Db     111 TVEWGEFQVDGRSSRGGQQTVNDAAAEALNAPDKRTALQIIKEKLPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180


RESULT   4
Q9WHF6
ID   Q9WHF6      PRELIMINARY;      PRT;   226 AA.
AC   Q9WHF6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   tomato mild mottle geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX   NCBI_TaxID=92943;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HN96-H5KW;
RA   Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA   Maxwell D.P.;
RT   "Molecular characterization and DNA-based detection methods for
RT   vegetable-infecting geminiviruses in Central America.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF131071; AAD33471.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     226    226
SQ   SEQUENCE   226 AA;  25941 MW;  2EA4116712871A23 CRC64;


  Query Match             76.2%;  Score 272;  DB 12;  Length 226;
  Best Local Similarity   72.9%;  Pred. No. 2.3e-23;
  Matches   51;  Conservative   12;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: |||||:|||||||| ||:|||||||||||||||||::|||:||:|:::|| :  |:|||
Db     111 TIEWGEFQIDGRSARGGQQTANDAAAEALNASSKEEAMRIIKEKLPKEFLFQYHNLSSNL 170

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180


RESULT   5
Q67574
ID   Q67574      PRELIMINARY;      PRT;   361 AA.
AC   Q67574;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE  REPLICATIVE PROTEIN.
GN   AL1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,
RA   Maxwell D.P., Russell D.R.;
RT   "Cloning of the complete DNA genomes of four bean-infecting
RT   geminiviruses and determining their infectivity by electric discharge
RT   particle acceleration.";
RL   Phytopathology 81:980-985(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,
RA   Morales F.J., Maxwell D.P.;
RT   "Differentiation of bean-infecting geminiviruses by nucleic acid
```

```
RT   hybridization probes and aspects of bean golden mosaic in Brazil.";
RL   Plant Dis. 75:336-342(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RT   "Genetic diversity in geminiviruses causing bean golden mosaic
RT   disease: The nucleotide sequence of the infectious cloned DNA
RT   components of a Brazilian isolate of bean golden mosaic geminivirus.";
RL   Submitted (MAY-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; M88686; AAA46312.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   361 AA;  41041 MW;  0094C7ACAF06B788 CRC64;


  Query Match             75.6%;  Score 270;  DB 12;  Length 361;
  Best Local Similarity   76.1%;  Pred. No. 6.8e-23;
  Matches   51;  Conservative    9;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLDRI 63
          || |||||||||| ||:|||:||||||||||:|||:||:|||:|| :  |:||||||
Db    113 WGHFQVDGRSARGGQQTANDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy     64 FDKTPEP 70
          | | |:|
Db    173 FTKAPDP 179


RESULT   6
Q9QDB1
ID   Q9QDB1      PRELIMINARY;      PRT;   225 AA.
AC   Q9QDB1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   cowpea golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=69263;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CGMV-BR;
RA   Faria J.C.;
RT   "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT   Brazil.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF188708; AAF06318.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     225    225
SQ   SEQUENCE   225 AA;  25766 MW;  1089CB6BD8D15B5D CRC64;


  Query Match             75.1%;  Score 268;  DB 12;  Length 225;
  Best Local Similarity   76.1%;  Pred. No. 6.6e-23;
  Matches   51;  Conservative    8;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLDRI 63
          || ||:|||||||| || ||||:||||||||||:|||:||:|||:|| :  |:||||||
Db    113 WGHFQIDGRSARGGQQTINDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy     64 FDKTPEP 70
          | | |||
Db    173 FKKPPEP 179


RESULT   7
Q9ELT8
ID   Q9ELT8      PRELIMINARY;      PRT;   314 AA.
AC   Q9ELT8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION ASSOCIATION PROTEIN.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF288227; AAG01006.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   314 AA;  35153 MW;  6B6220613046943F CRC64;


  Query Match             75.1%;  Score 268;  DB 12;  Length 314;
  Best Local Similarity   79.4%;  Pred. No. 9.8e-23;
  Matches   54;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
          |: |||||||||||||| ||:|||||||||| ||| |||||||:|||||| |  | |||
Db    110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy     61 DRIFDKTP 68
          ||||   |
Db    170 DRIFSPPP 177


RESULT   8
Q9QS55
ID   Q9QS55      PRELIMINARY;      PRT;   364 AA.
AC   Q9QS55;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION INITIATION PROTEIN AC1.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A.;
RT   "Cloning of a DNA-A-like genomic component of sweet potato leaf curl
RT   virus : nucleotide sequence and phylogenetic relationships.";
RL   Online Publication.
DR   EMBL; AF104036; AAD47173.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   364 AA;  40680 MW;  5F79752431A09D6E CRC64;
```

```
  Query Match             75.1%;  Score 268;  DB 12;  Length 364;
  Best Local Similarity   79.4%;  Pred. No. 1.2e-22;
  Matches   54;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: |||||||||||||| ||:|||||||||| ||| |||||||||:|||||| |   | |||
Db     110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy      61 DRIFDKTP 68
           ||||   |
Db     170 DRIFSPPP 177


RESULT   9
Q98693
ID   Q98693      PRELIMINARY;      PRT;   185 AA.
AC   Q98693;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   AC1.
OS   sida golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51034;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA;
RA   Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RT   "Genetic Diversity among geminiviruses associated with the weed
RT   species Sida spp, Macroptilium lathyroides, and Wissadula amplissima
RT   from Jamaica.";
RL   Plant Dis. 81:1251-1258(1997).
DR   EMBL; U67926; AAB97865.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     185    185
SQ   SEQUENCE   185 AA;  20975 MW;  3913850A025A5EE1 CRC64;


  Query Match             74.2%;  Score 265;  DB 12;  Length 185;
  Best Local Similarity   70.0%;  Pred. No. 1.2e-22;
  Matches   49;  Conservative   12;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: || ||:|||||||| ||:|||||||||: :||:||:|||||:|||:||| :  |:||:
Db      89 TIEWGVFQIDGRSARGGQQTANDAAAEALNSGTKEDALKIIREKLPERYLFQYHNLSSNI 148

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     149 DRIFSKPPEP 158


RESULT  10
P88975
ID   P88975      PRELIMINARY;      PRT;   149 AA.
AC   P88975;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   AC1.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICAN;
RA   Roye M.E.;
RL   Thesis (1996), Biochemistry, University of the West Indies, Jamaica.
DR   EMBL; U75278; AAB36919.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     149    149
SQ   SEQUENCE   149 AA;  16785 MW;  E4CF5EED4C9CD508 CRC64;


  Query Match             72.0%;  Score 257;  DB 12;  Length 149;
  Best Local Similarity   68.6%;  Pred. No. 7.5e-22;
  Matches   48;  Conservative   12;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: || ||:|||||||| ||||||||||||: :|| |::|::||:|||:||  :  |:|||
Db      52 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 111

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     112 DRIFMKDPEP 121


RESULT  11
Q9YLA4
ID   Q9YLA4      PRELIMINARY;      PRT;   233 AA.
AC   Q9YLA4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E.;
RT   "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
RT   from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E., McLaughlin W.A., Maxwell D.P.;
RT   "Molecular characterization of two distinct geminiviruses infecting M.
RT   lathyroides from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF098940; AAD17850.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     233    233
SQ   SEQUENCE   233 AA;  26356 MW;  AA490AF4D2166A02 CRC64;


  Query Match             72.0%;  Score 257;  DB 12;  Length 233;
  Best Local Similarity   68.6%;  Pred. No. 1.3e-21;
  Matches   48;  Conservative   12;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: || ||:|||||||| ||||||||||||: :|| |::|::||:|||:||  :  |:|||
Db     110 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 169

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     170 DRIFMKDPEP 179


RESULT  12
O39180
```

```
ID   O39180      PRELIMINARY;      PRT;   234 AA.
AC   O39180;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   AC1 (FRAGMENT).
OS   Geminiviridae.
OC   Viruses; ssDNA viruses.
OX   NCBI_TaxID=10811;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOMATO STRAIN;
RA   Guzman P., Arredondo C.R., Emmatty D., Portillo R.J., Gilbertson R.L.;
RT   "Partial Characterization of Two Whitefly-Transmitted Geminiviruses
RT   Infecting Tomatoes in Venezuela.";
RL   Plant Dis. 81:312-312(1997).
DR   EMBL; AF026553; AAB82605.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     234     234
SQ   SEQUENCE   234 AA;  26486 MW;  9ED8F0697105CD19 CRC64;


  Query Match             70.3%;  Score 251;  DB 12;  Length 234;
  Best Local Similarity   67.1%;  Pred. No. 6.2e-21;
  Matches   47;  Conservative   13;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: || ||:|||||||| ||:|||||||||: :||||::||:||:|||:|| :  |:|:|
Db     110 TIEWGLFQIDGRSARGGQQTANDAAAEALNSGTKEEAMKIIKEKLPEKFLFQYHNLSSDL 169

Qy      61 DRIFDKTPEP 70
           |||| | | |
Db     170 DRIFMKAPNP 179


RESULT  13
Q9Z089
ID   Q9Z089      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z089;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA3-1, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001315; BAA34033.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1       1
FT   NON_TER     190     190
SQ   SEQUENCE   190 AA;  21432 MW;  AAC093D1D1610FAD CRC64;


  Query Match             69.5%;  Score 248;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 1.1e-20;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           || || ||:|||||||||| :||| |||||||| ||| |||||:|: ::| :  |||||
Db      85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  14
Q9Z084
ID   Q9Z084      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z084;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA5-2, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA5;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001318; BAA34039.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1       1
FT   NON_TER     190     190
SQ   SEQUENCE   190 AA;  21444 MW;  AAC1C2943E3F01AD CRC64;


  Query Match             69.5%;  Score 248;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 1.1e-20;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           || || ||:|||||||||| :||| |||||||| ||| |||||:|: ::| :  |||||
Db      85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  15
Q9W827
ID   Q9W827      PRELIMINARY;      PRT;   190 AA.
AC   Q9W827;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE GORA1-1, PARTIAL AND
DE   COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GORA1;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001303; BAA34010.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
```

```
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  93C3742A8EBDB7EB CRC64;


  Query Match             69.2%;  Score 247;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 1.4e-20;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           || || |||||||||||||| :||| |||||||:|| ||| |||||:|: ::| :  |||||
Db      85 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFSNSSFDQVPE 169


Search completed: October 10, 2002, 02:01:59
Job time: 1092 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:09:52 ; Search time 34.62 Seconds
(without alignments)
78.289 Million cell updates/sec

Title: US-09-289-346A-6
Perfect score: 357
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FAFAALNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	340	95.2	352	1	VAL1_TGMV	P03567 tomato gold
2	240	67.2	361	1	VAL1_PYMVV	P27258 potato yell
3	227	63.6	358	1	VAL1_CLVK	P14982 cassava lat
4	227	63.6	358	1	VAL1_CLVN	P14972 cassava lat
5	220	61.6	362	1	VAL1_TYLCA	P36279 tomato yell
6	217	60.8	359	1	VAL1_TYLCU	P38609 tomato yell
7	210	58.8	359	1	VAL1_TYLCM	P27260 tomato yell
8	209	58.5	349	1	VAL1_PHUV	Q06923 pepper huas
9	204	57.1	353	1	VAL1_BGMV	P05175 bean golden
10	203	56.9	355	1	VAL1_ABMVW	P21947 abutilon mo
11	200	56.0	358	1	VAL1_BCTV	P14991 beet curly
12	200	56.0	361	1	VAL1_TMOV	Q06657 tomato mott
13	191	53.5	357	1	VAL1_TYLCV	P27259 tomato yell
14	104	29.1	347	1	VAL1_SLCV	P29048 squash leaf
15	67	18.8	447	1	TBL2_HUMAN	Q9y4p3 homo sapien
16	64.5	18.1	1713	1	LMA3_HUMAN	Q16787 homo sapien
17	63.5	17.8	1610	1	CCAD_MESAU	Q99244 mesocricetu
18	63.5	17.8	2161	1	CCAD_HUMAN	Q01668 homo sapien
19	63.5	17.8	2203	1	CCAD_RAT	P27732 rattus norv
20	60	16.8	269	1	ORN_YEAST	P54964 saccharomyc
21	60	16.8	295	1	VAL1_TYDVA	P31617 tobacco yel
22	60	16.8	355	1	CRTB_RHOSH	P54905 rhodobacter
23	59.5	16.7	299	1	Y175_HELPJ	Q9zmq7 helicobacte
24	59.5	16.7	470	1	RHBA_RHIME	Q9z3r2 rhizobium m
25	59.5	16.7	4385	1	YP73_CAEEL	Q09222 caenorhabdi
26	58.5	16.4	706	1	Y006_RICPR	Q9zed6 rickettsia
27	58.5	16.4	1232	1	YOQ5_CAEEL	P34643 caenorhabdi
28	58	16.2	207	1	IL6_MARMO	O35736 marmota mon
29	58	16.2	319	1	LDH2_STAAM	Q99r35 staphylococ
30	58	16.2	447	1	CDSA_DROME	P56079 d phosphati
31	57.5	16.1	367	1	LHX4_MOUSE	P53776 mus musculu
32	57.5	16.1	844	1	SECA_STACA	P47994 staphylococ
33	57.5	16.1	1852	1	CCAS_CYPCA	P22316 cyprinus ca
34	57	16.0	1483	1	CYP1_YEAST	P12351 saccharomyc
35	57	16.0	1608	1	HLYA_SERMA	P15320 serratia ma
36	56.5	15.8	511	1	HUTH_VIBCH	Q9ksq4 vibrio chol
37	56.5	15.8	2190	1	CCAD_CHICK	O73700 gallus gall
38	56	15.7	258	1	YL48_MYCTU	O06228 mycobacteri
39	56	15.7	408	1	CINA_THEMA	Q9s5x1 thermotoga
40	56	15.7	492	1	MOT3_MOUSE	O35308 mus musculu
41	56	15.7	492	1	MOT3_RAT	O70461 rattus norv
42	56	15.7	703	1	GYS2_RAT	P17625 rattus norv
43	56	15.7	767	1	TOP1_MOUSE	Q04750 mus musculu
44	56	15.7	1033	1	YDK9_SCHPO	P87115 schizosacch
45	56	15.7	1051	1	CARB_SULSO	Q59969 sulfolobus

ALIGNMENTS

RESULT 1
VAL1_TGMV
ID VAL1_TGMV STANDARD; PRT; 352 AA.
AC P03567;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein.
GN AC1.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10831;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;
RT "Complete nucleotide sequence of the infectious cloned DNA components
RT of tomato golden mosaic virus: potential coding regions and regulatory
RT sequences.";
RL EMBO J. 3:2197-2205(1984).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; K02029; -; NOT_ANNOTATED_CDS.
DR PIR; A04170; QQCVL1.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR ProDom; PD000736; Gemini_AL1; 1.
KW ATP-binding.
FT NP_BIND 223 230 ATP (POTENTIAL).
SQ SEQUENCE 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;

Query Match 95.2%; Score 340; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 1.2e-30;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |||||||||||||||||||||||||||||||||||||||||||||||||| |  |||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
VAL1_PYMVV
ID VAL1_PYMVV STANDARD; PRT; 361 AA.

```
AC   P27258;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 protein.
OS   Potato yellow mosaic virus (isolate Venezuela).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10828;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91311403; PubMed=1856690;
RA   Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT   "The nucleotide sequence of the infectious cloned DNA components of
RT   potato yellow mosaic virus.";
RL   J. Gen. Virol. 72:1515-1520(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00940; BAA00782.1; -.
DR   PIR; JU0364; QQCVPT.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   361 AA;  40850 MW;  5627A33BF1264383 CRC64;


  Query Match             67.2%;  Score 240;  DB 1;  Length 361;
  Best Local Similarity   66.7%;  Pred. No. 1.7e-19;
  Matches   46;  Conservative   11;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: || ||:|||||||| || |||||||||: :|| |::||:||:|||:|| :  |: ||
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 DRIFDKTPE 69
           |||| | ||
Db     170 DRIFMKAPE 178


RESULT   3
VAL1_CLVK
ID   VAL1_CLVK      STANDARD;      PRT;   358 AA.
AC   P14982;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (40.4 kDa protein).
GN   AC1.
OS   Cassava latent virus (strain West Kenyan 844).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10818;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Gay M.R.;
RT   "Nucleotide sequence of cassava latent virus DNA.";
RL   Nature 301:260-262(1983).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J02057; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40346 MW;  ED173E753EE92D69 CRC64;


  Query Match             63.6%;  Score 227;  DB 1;  Length 358;
  Best Local Similarity   60.0%;  Pred. No. 4.8e-18;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: ||:||:|||||||| |::||| |:|||: || ||| :||| :|: ::  |  |||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   4
VAL1_CLVN
ID   VAL1_CLVN      STANDARD;      PRT;   358 AA.
AC   P14972;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (40.4 kDa protein).
GN   AC1.
OS   Cassava latent virus (strain Nigerian).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10819;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90174930; PubMed=2308831;
RA   Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT   "Nucleotide sequence of the infectious cloned DNA components of
RT   African cassava mosaic virus (Nigerian strain).";
RL   Nucleic Acids Res. 18:197-198(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X17095; CAA34953.1; -.
DR   PIR; S07594; S07594.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40435 MW;  1DB16BB0CB2D5E2C CRC64;


  Query Match             63.6%;  Score 227;  DB 1;  Length 358;
  Best Local Similarity   60.0%;  Pred. No. 4.8e-18;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: ||:||:|||||||| |::||| |:|||: || ||| :||| :|: ::  |  |||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168
```

```
Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   5
VAL1_TYLCA
ID   VAL1_TYLCA     STANDARD;      PRT;   362 AA.
AC   P36279;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=36447;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93139778; PubMed=8423446;
RA   Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;
RT   "Nucleotide sequence and genome organization of tomato leaf curl
RT   geminivirus.";
RL   J. Gen. Virol. 74:147-151(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
DR   PIR; JQ1887; JQ1887.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   362 AA;  41197 MW;  343E7184B4704098 CRC64;


  Query Match             61.6%;  Score 220;  DB 1;  Length 362;
  Best Local Similarity   52.9%;  Pred. No. 3e-17;
  Matches   45;  Conservative    9;  Mismatches   15;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           || ||||:|||||||| |::||| |:||| || ||| ::|| |: |:  |  |||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   6
VAL1_TYLCU
ID   VAL1_TYLCU     STANDARD;      PRT;   359 AA.
AC   P38609;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=37139;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94256836; PubMed=8198442;
RA   Noris E., Hidalgo E., Accotto G., Moriones E.;
RT   "High similarity among the tomato yellow leaf curl virus isolates
RT   from the west Mediterranean basin: the nucleotide sequence of an
RT   infectious clone from Spain.";
RL   Arch. Virol. 135:165-170(1994).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z25751; CAA81026.1; -.
DR   PIR; S39211; S39211.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  41065 MW;  2D170A51EF80A3EC CRC64;


  Query Match             60.8%;  Score 217;  DB 1;  Length 359;
  Best Local Similarity   59.4%;  Pred. No. 6.3e-17;
  Matches   41;  Conservative    9;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || ||| :|:|  |  |:  |  :|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   7
VAL1_TYLCM
ID   VAL1_TYLCM     STANDARD;      PRT;   359 AA.
AC   P27260;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10833;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92107660; PubMed=1840676;
RA   Kheyr-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,
RA   Gronenborn B.;
RT   "Tomato yellow leaf curl virus from Sardinia is a
RT   whitefly-transmitted monopartite geminivirus.";
RL   Nucleic Acids Res. 19:6763-6769(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X61153; CAA43466.1; -.
DR   PIR; S22593; S22593.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  40733 MW;  9717B4A07C93EFA7 CRC64;


  Query Match             58.8%;  Score 210;  DB 1;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 3.8e-16;
```

```
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |  :|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   8
VAL1_PHUV
ID   VAL1_PHUV      STANDARD;      PRT;   349 AA.
AC   Q06923;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AL1.
OS   Pepper huasteco virus (PHV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=28349;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94015007; PubMed=8409944;
RA   Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA   Rivera-Bustamante R.F.;
RT   "Complete nucleotide sequence of pepper huasteco virus: analysis and
RT   comparison with bipartite geminiviruses.";
RL   J. Gen. Virol. 74:2225-2231(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70418; CAA49856.1; -.
DR   PIR; S31875; S31875.
DR   PIR; JQ2300; JQ2300.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (BY SIMILARITY).
SQ   SEQUENCE   349 AA;  39722 MW;  D5F4E76CD56370F4 CRC64;


  Query Match             58.5%;  Score 209;  DB 1;  Length 349;
  Best Local Similarity   57.1%;  Pred. No. 4.8e-16;
  Matches   40;  Conservative   13;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: |||||:|||||||| |::||  |:|||::| |||||||:|: |: :   |  : ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   9
VAL1_BGMV
ID   VAL1_BGMV      STANDARD;      PRT;   353 AA.
AC   P05175;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   AL1 protein (40.2 kDa protein).
GN   AC1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Howarth A.J., Caton J., Bossert M., Goodman R.M.;
RT   "Nucleotide sequence of bean golden mosaic virus and a model for gene
RT   regulation in geminiviruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M10070; AAA46318.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   353 AA;  40190 MW;  80FA779DF6029A34 CRC64;


  Query Match             57.1%;  Score 204;  DB 1;  Length 353;
  Best Local Similarity   57.1%;  Pred. No. 1.8e-15;
  Matches   40;  Conservative   13;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: ||:||||||||||| |::||: |:|||| | | || |::|: |: |:     : |||
Db     110 TIEWGQFQVDGRSARGGQQSANDSYAKALNADSIESALTILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFFKVPEP 179


RESULT  10
VAL1_ABMVW
ID   VAL1_ABMVW     STANDARD;      PRT;   355 AA.
AC   P21947;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AC1.
OS   Abutilon mosaic virus (isolate West India).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10816;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91020984; PubMed=2219703;
RA   Frischmuth T., Zimmat G., Jeske H.;
RT   "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT   as well as eukaryotic features.";
RL   Virology 178:461-468(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR   PIR; A36214; QQCVW1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   355 AA;  40257 MW;  16A2CA8A63251E95 CRC64;


  Query Match             56.9%;  Score 203;  DB 1;  Length 355;
  Best Local Similarity   55.7%;  Pred. No. 2.3e-15;
  Matches   39;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |  |||||:||||||||  ||:||: |:||||   : || |::|: |: |:     : |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
VAL1_BCTV
ID   VAL1_BCTV      STANDARD;      PRT;   358 AA.
AC   P14991;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 protein (40.8 kDa protein).
OS   Beet curly top virus (BCTV).
OC   Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
OX   NCBI_TaxID=10840;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT   "The nucleotide sequence of an infectious clone of the geminivirus
RT   beet curly top virus.";
RL   EMBO J. 5:1761-1767(1986).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40889 MW;  39A45FE3C0B9C333 CRC64;


  Query Match             56.0%;  Score 200;  DB 1;  Length 358;
  Best Local Similarity   52.9%;  Pred. No. 5e-15;
  Matches   37;  Conservative   16;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: |||||:|||||||| ||:||: |:|||||:| ::||||::|: |: |      | :|
Db     110 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 169

Qy      61 DRIFDKTPEP 70
           :|| : |:|
Db     170 QKIFQRPPDP 179


RESULT  12
VAL1_TMOV
ID   VAL1_TMOV      STANDARD;      PRT;   361 AA.
AC   Q06657;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AL1.
OS   Tomato mottle virus (isolate Florida) (TMoV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=36449;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93107858; PubMed=1469361;
RA   Abouzid A.M., Polston J.E., Hiebert E.;
RT   "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT   isolated from tomatoes in Florida.";
RL   J. Gen. Virol. 73:3225-3229(1992).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L14460; AAC32414.1; -.
DR   PIR; JQ1870; JQ1870.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (BY SIMILARITY).
SQ   SEQUENCE   361 AA;  40516 MW;  8138B65CEEAC6950 CRC64;


  Query Match             56.0%;  Score 200;  DB 1;  Length 361;
  Best Local Similarity   52.9%;  Pred. No. 5e-15;
  Matches   37;  Conservative   16;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: ||:||:|||||||| |::||: |:|||| | : || ::||: |: ::     : |||
Db     110 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  13
VAL1_TYLCV
ID   VAL1_TYLCV      STANDARD;      PRT;   357 AA.
AC   P27259;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10832;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92024070; PubMed=1926771;
RA   Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;
RT   "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RT   with a single genomic component.";
RL   Virology 185:151-161(1991).
```

```
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15656; CAA33688.1; -.
DR   PIR; D40779; QQCVC1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     219    226       ATP (POTENTIAL).
SQ   SEQUENCE   357 AA;  40678 MW;  939AB68E1AB3B2A7 CRC64;


  Query Match             53.5%;  Score 191;  DB 1;  Length 357;
  Best Local Similarity   62.3%;  Pred. No. 5e-14;
  Matches   38;  Conservative   10;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLDRI 63
           :|  |:||||||||| |::||| |||||: || ||| |::|| |: |:  |  |:||||||
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171


RESULT  14
VAL1_SLCV
ID   VAL1_SLCV      STANDARD;      PRT;   347 AA.
AC   P29048;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   AL1 protein.
OS   Squash leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10829;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91082449; PubMed=1984668;
RA   Lazarowitz S.G., Lazdins I.B.;
RT   "Infectivity and complete nucleotide sequence of the cloned genomic
RT   components of a bipartite squash leaf curl geminivirus with a broad
RT   host range phenotype.";
RL   Virology 180:58-69(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M38183; AAC32410.1; ALT_INIT.
DR   PIR; C36785; QQCVS1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     218    225       ATP (POTENTIAL).
SQ   SEQUENCE   347 AA;  39110 MW;  AFDAEBDDE122110E CRC64;


  Query Match             29.1%;  Score 104;  DB 1;  Length 347;
  Best Local Similarity   34.8%;  Pred. No. 0.00024;
  Matches   23;  Conservative   14;  Mismatches   25;  Indels    4;  Gaps    1;

Qy       5 GEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLDRIF 64
           |:::| |    |     :|    |:|| |  ||| ||:   |: ::  :  | :|::|:|
Db     116 GQYKVSG----GSKSNKDDVYHNAVNAGSAGEALDIIKAGDPKTFIVNYHNLLANVERLF 171

Qy      65 DKTPEP 70
             | |||
Db     172 QKPPEP 177


RESULT  15
TBL2_HUMAN
ID   TBL2_HUMAN     STANDARD;      PRT;   447 AA.
AC   Q9Y4P3; Q9UQE2;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Transducin beta-like 2 protein (WS beta-transducin repeats protein)
DE   (WS-betaTRP).
GN   TBL2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20044626; PubMed=10575226;
RA   Perez Jurado L.A., Wang Y.-K., Francke U., Cruces J.;
RT   "TBL2, a novel transducin family member in the WBS deletion:
RT   characterization of the complete sequence, genomic structure,
RT   transcriptional variants and the mouse ortholog.";
RL   Cytogenet. Cell Genet. 86:277-284(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=21154917; PubMed=11230166;
RA   Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA   Ansorge W., Boecher M., Bloecker H., Bauersachs S., Blum H.,
RA   Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
RA   Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA   Wambutt R., Korn B., Klein M., Poustka A.;
RT   "Towards a catalog of human genes and proteins: sequencing and
RT   analysis of 500 novel complete protein coding human cDNAs.";
RL   Genome Res. 11:422-435(2001).
RN   [3]
RP   SEQUENCE OF 14-447 FROM N.A.
RX   MEDLINE=99075645; PubMed=9860302;
RA   Meng X., Lu X., Li Z., Green E.D., Massa H., Trask B.J., Morris C.A.,
RA   Keating M.T.;
RT   "Complete physical map of the common deletion region in Williams
RT   syndrome and identification and characterization of three novel
RT   genes.";
RL   Hum. Genet. 103:590-599(1998).
CC   -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF097484; AAF06823.1; -.
DR   EMBL; AF097485; AAF06824.1; -.
DR   EMBL; AL080162; CAB45751.1; -.
DR   EMBL; AF056183; AAD28083.1; ALT_INIT.
DR   MIM; 605842; -.
DR   InterPro; IPR001680; WD40.
```

```
DR   Pfam; PF00400; WD40; 5.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   SMART; SM00320; WD40; 5.
DR   PROSITE; PS00678; WD_REPEATS_1; 1.
DR   PROSITE; PS50082; WD_REPEATS_2; 3.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Repeat; WD repeat.
FT   REPEAT       88    127       WD 1.
FT   REPEAT      134    174       WD 2.
FT   REPEAT      186    226       WD 3.
FT   REPEAT      228    267       WD 4.
FT   REPEAT      277    316       WD 5.
FT   REPEAT      329    367       WD 6.
FT   REPEAT      371    409       WD 7.
SQ   SEQUENCE   447 AA;  49797 MW;  B260087E1A71D3F9 CRC64;


  Query Match           18.8%;  Score 67;  DB 1;  Length 447;
  Best Local Similarity  32.1%;  Pred. No. 4.3;
  Matches   18;  Conservative   11;  Mismatches   23;  Indels    4;  Gaps    1;

Qy      12 RSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAF----AALNSNLDRI 63
           || |  || :|    :  : | |::  | ||:: |::: |      ||| |:   |
Db      38 RSGRPACQKANGFPPDKSSGSKKQKQYQRIRKEKPQQHNFTHRLLAAALKSHSGNI 93


Search completed: October 10, 2002, 02:09:52
Job time: 1350 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:48:44 ; Search time 69.63 Seconds
(without alignments)
96.600 Million cell updates/sec

Title: US-09-289-346A-6
Perfect score: 357
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FAFAALNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	340	95.2	352	1	QQCVL1	AL1 protein - toma
2	240	67.2	361	1	QQCVPT	AL1 protein - pota
3	227	63.6	358	2	S07594	hypothetical prote
4	220	61.6	362	1	JQ1887	AL1 protein - toma
5	217	60.8	359	2	S39211	gene C1 protein -
6	212	59.4	351	2	JQ2327	AL1 protein - Indi
7	210	58.8	359	2	S22593	hypothetical prote
8	209	58.5	349	2	JQ2300	replicase - pepper
9	209	58.5	349	2	S31875	AL1 protein - pepp
10	204	57.1	359	2	S39235	gene C1 protein -
11	203	56.9	355	1	QQCVW1	AV1 protein - abut
12	200	56.0	358	1	JQ1870	AL1 protein - toma
13	200	56.0	385	2	S28360	AL1 protein - beet
14	193	54.1	360	2	S59885	replication-associ
15	191	53.5	357	1	QQCVC1	AL1 protein - toma
16	118	33.1	131	2	S45059	AC1 protein (clone
17	104	29.1	347	1	QQCVS1	AL1 protein - squa
18	67	18.8	447	2	T12544	hypothetical prote
19	64.5	18.1	587	2	JC1419	Fc gamma (IgG) rec
20	64.5	18.1	1713	2	A55347	adhesive ligand ep
21	63.5	17.8	180	2	D84082	hypothetical prote
22	63.5	17.8	1610	2	A46227	voltage-dependent
23	63.5	17.8	1646	2	JH0422	voltage-dependent
24	63.5	17.8	2161	2	JH0564	calcium channel al
25	63.5	17.8	2181	2	A38198	calcium channel al
26	63.5	17.8	2203	2	T42742	voltage-dependent
27	63	17.6	316	2	C82085	conserved hypothet
28	62.5	17.5	1033	2	E97700	hypothetical prote
29	62	17.4	338	2	AG3582	iron(III)-binding
30	62	17.4	340	2	B97011	probably O-sialogl
31	61.5	17.2	392	2	T45290	acetyl-CoA C-acety
32	60.5	16.9	840	2	T36175	probable large ATP
33	60	16.8	154	2	AD3475	nitrogen regulator
34	60	16.8	269	2	S61633	YNT20 protein - ye
35	60	16.8	295	2	D42452	C1 protein - tobac
36	60	16.8	355	2	S49621	phytoene synthase
37	60	16.8	355	2	T50746	phytoene synthase
38	60	16.8	714	2	C95382	probable ferrichro
39	60	16.8	827	2	D96907	phage-related prot
40	60	16.8	1888	2	T14273	zinc finger protei
41	59.5	16.7	299	2	B71967	probable peptidyl-
42	59.5	16.7	470	2	T46814	gamma-aminobutyrat
43	59.5	16.7	470	2	B95419	diaminobutyrate--p
44	59.5	16.7	2137	2	T05244	hypothetical prote
45	59.5	16.7	4385	2	T29042	hypothetical prote

ALIGNMENTS

RESULT 1
QQCVL1
AL1 protein - tomato golden mosaic virus
C;Species: tomato golden mosaic virus
A;Note: host Nicotiana sp. (tobacco)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C;Accession: A04170
R;Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.
EMBO J. 3, 2197-2205, 1984
A;Title: Complete nucleotide sequence of the infectious cloned DNA components of toma
A;Reference number: A04163
A;Accession: A04170
A;Molecule type: DNA
A;Residues: 1-352 <HAM>
C;Comment: The genome consists of two circular, single-stranded DNA components, DNA A
.
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 95.2%; Score 340; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 4.6e-31;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
QQCVPT
AL1 protein - potato yellow mosaic virus (isolate Venezuela)
C;Species: potato yellow mosaic virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C;Accession: JU0364
R;Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.
J. Gen. Virol. 72, 1515-1520, 1991
A;Title: The nucleotide sequence of the infectious cloned DNA components of potato ye
A;Reference number: JU0362; MUID:91311403
A;Accession: JU0364
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-361 <COU>
A;Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             67.2%;  Score 240;  DB 1;  Length 361;
  Best Local Similarity   66.7%;  Pred. No. 1.2e-19;
  Matches   46;  Conservative   11;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: || ||:||||||||| || |||||||||: :|| |::||:||:|||:|| :  |: ||
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 DRIFDKTPE 69
           |||| | ||
Db     170 DRIFMKAPE 178
```

RESULT 3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C;Species: cassava latent virus
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C;Accession: S07594
R;Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A;Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A;Reference number: S07590; MUID:90174930
A;Accession: S07594
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <MOR>
A;Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C;Genetics:
A;Map position: segment DNA1
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             63.6%;  Score 227;  DB 2;  Length 358;
  Best Local Similarity   60.0%;  Pred. No. 3.7e-18;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: ||:||:|||||||| |::||| |:|||: || ||| :||| :|: ::  |  |||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178
```

RESULT 4
JQ1887
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1887
R;Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A;Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A;Reference number: JQ1885; MUID:93139778
A;Accession: JQ1887
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-362 <DRY>
A;Cross-references: GB:S53251
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             61.6%;  Score 220;  DB 1;  Length 362;
  Best Local Similarity   52.9%;  Pred. No. 2.4e-17;
  Matches   45;  Conservative    9;  Mismatches   15;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           || |||||:|||||||| |::||| |:|||  || ||| ::||  |: |:  |  |||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194
```

RESULT 5
S39211
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C;Accession: S39211
R;Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A;Description: High similarity among the tomato yellow leaf curl virus isolates from
A;Reference number: S39209
A;Accession: S39211
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <NOR>
A;Cross-references: EMBL:Z25751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             60.8%;  Score 217;  DB 2;  Length 359;
  Best Local Similarity   59.4%;  Pred. No. 5.1e-17;
  Matches   41;  Conservative    9;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || ||| :|:|  |  |:  |  :||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179
```

RESULT 6
JQ2327
AL1 protein - Indian cassava mosaic virus
N;Alternate names: replication-associated protein
C;Species: Indian cassava mosaic virus
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C;Accession: JQ2327; S35883
R;Hong, Y.G.; Robinson, D.J.; Harrison, B.D.
J. Gen. Virol. 74, 2437-2443, 1993
A;Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-t
A;Reference number: JQ2326; MUID:94065670
A;Accession: JQ2327
A;Molecule type: DNA
A;Residues: 1-351 <HON>
A;Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             59.4%;  Score 212;  DB 2;  Length 351;
  Best Local Similarity   62.7%;  Pred. No. 1.9e-16;
  Matches   42;  Conservative    8;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLDRI 63
           || ||:|||||||| |::||| | |||: || |||:|:||  |  ||  |  ::||||||
Db     113 WGTFQIDGRSARGGQQSANDAYAAALNSGSKSEALKILRELAPRDYLRDFHHISSNLDRI 172

Qy      64 FDKTPEP 70
           | | | |
Db     173 FTKPPPP 179
```

RESULT 7
S22593
hypothetical protein C4 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus

```
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Sep-1999
C;Accession: S22593
R;Kheyr-Pour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.
Nucleic Acids Res. 19, 6763-6769, 1991
A;Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monoparti
A;Reference number: S22588; MUID:92107660
A;Accession: S22593
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-359 <KHE>
A;Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             58.8%;  Score 210;  DB 2;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 3.2e-16;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLD 61
          | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |  :|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   8
JQ2300
replicase - pepper huasteco virus (component A)
N;Alternate names: ORF AL1 protein
C;Species: pepper huasteco virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C;Accession: JQ2300
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante, R.F
J. Gen. Virol. 74, 2225-2231, 1993
A;Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with bipa
A;Reference number: JQ2299; MUID:94015007
A;Accession: JQ2300
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             58.5%;  Score 209;  DB 2;  Length 349;
  Best Local Similarity   57.1%;  Pred. No. 4.1e-16;
  Matches   40;  Conservative   13;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: ||||:|||||||| |::|| |:|||::| ||||||:|: |: :   |  : ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   9
S31875
AL1 protein - pepper rizado amarillo virus
C;Species: pepper rizado amarillo virus
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Sep-1999
C;Accession: S31875
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera-Bu
submitted to the EMBL Data Library, February 1993
A;Description: Complete nucleotide sequence of pepper huasteco virus: analysis and compa
A;Reference number: S31872
A;Accession: S31875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
A;Note: the source is designated as pepper huasteco virus
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             58.5%;  Score 209;  DB 2;  Length 349;
  Best Local Similarity   57.1%;  Pred. No. 4.1e-16;
  Matches   40;  Conservative   13;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: ||||:|||||||| |::|| |:|||::| ||||||:|: |: :   |  : ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT  10
S39235
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Sep-1997
C;Accession: S39235
R;Crespi, S.; Noris, E.; Vaira, A.; Bosco, D.; Accotto, G.
submitted to the EMBL Data Library, December 1993
A;Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.
A;Reference number: S39233
A;Accession: S39235
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <CRE>
A;Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             57.1%;  Score 204;  DB 2;  Length 359;
  Best Local Similarity   55.1%;  Pred. No. 1.6e-15;
  Matches   38;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLD 61
          | || ||:|||||||| ||:||| |:|:|| || ||| :|::  |  |:  |  ::||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINARSKSEALDVIKQLAPRDYVLHFHNISSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT  11
QQCVW1
AV1 protein - abutilon mosaic virus (isolate West India)
C;Species: abutilon mosaic virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Apr-1994
C;Accession: A36214
R;Frischmuth, T.; Zimmat, G.; Jeske, H.
Virology 178, 461-468, 1990
A;Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as
A;Reference number: A36214; MUID:91020984
A;Accession: A36214
A;Molecule type: DNA
A;Residues: 1-355 <FRI>
A;Cross-references: EMBL:X15983
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             56.9%;  Score 203;  DB 1;  Length 355;
  Best Local Similarity   55.7%;  Pred. No. 2e-15;
  Matches   39;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
```

```
          |  |||||:||||||||  ||:||:  |:||||   : || |::|:  |: |:     : |||
Db    110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 ERIFAKAPEP 179


RESULT  12
JQ1870
AL1 protein - tomato mottle virus (isolate Florida)
C;Species: tomato mottle virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1870
R;Abouzid, A.M.; Polston, J.E.; Hiebert, E.
J. Gen. Virol. 73, 3225-3229, 1992
A;Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated from
A;Reference number: JQ1869; MUID:93107858
A;Accession: JQ1870
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <ABO>
A;Cross-references: GB:L14460
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             56.0%;  Score 200;  DB 1;  Length 358;
  Best Local Similarity   52.9%;  Pred. No. 4.5e-15;
  Matches   37;  Conservative   16;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
          |: ||:||:|||||||| |::||: |:|||| | : || ::||: |: ::     : |||
Db    107 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 166

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    167 ERIFAKAPEP 176


RESULT  13
S28360
AL1 protein - beet curly top virus
C;Species: beet curly top virus
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
C;Accession: S28360
R;Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
EMBO J. 5, 1761-1767, 1986
A;Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly to
A;Reference number: S28360
A;Accession: S28360
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-385 <STA>
A;Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             56.0%;  Score 200;  DB 2;  Length 385;
  Best Local Similarity   52.9%;  Pred. No. 4.8e-15;
  Matches   37;  Conservative   16;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
          |: |||||:|||||||| ||:||: |:||||:| ::||||::|: |: |      | :|
Db    137 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 196

Qy     61 DRIFDKTPEP 70
          :|| : |:|
Db    197 QKIFQRPPDP 206


RESULT  14
S59885
replication-associated protein C1 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
C;Accession: S59885
R;Hong, Y.; Harrison, B.D.
submitted to the EMBL Data Library, February 1995
A;Description: Nucleotide sequences from tomato leaf curl viruses from different coun
d geminiviruses.
A;Reference number: S58346
A;Accession: S59885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <HON>
A;Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA88229.1; PID:g974211
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             54.1%;  Score 193;  DB 2;  Length 360;
  Best Local Similarity   57.6%;  Pred. No. 2.8e-14;
  Matches   38;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLDRI 63
          :| ||:|||||||| |::||| |||:|: ||  || |:||| |: ::  |  ||:|||||
Db    113 FGVFQIDGRSARGGQQSANDAYAEAINSGSKAAALDILREKAPKDFVLQFHNLNANLDRI 172

Qy     64 FDKTPE 69
          |  : |
Db    173 FTPSAE 178


RESULT  15
QQCVC1
AL1 protein - tomato yellow leaf curl virus
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C;Accession: D40779
R;Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.
Virology 185, 151-161, 1991
A;Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin
A;Reference number: A40779; MUID:92024070
A;Accession: D40779
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-357 <NAV>
A;Cross-references: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             53.5%;  Score 191;  DB 1;  Length 357;
  Best Local Similarity   62.3%;  Pred. No. 4.7e-14;
  Matches   38;  Conservative   10;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLDRI 63
          :|  |:|||||||| |::||| |||||: || ||| |::|| |: |:  |  |:||||||
Db    111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy     64 F 64
          |
Db    171 F 171


Search completed: October 10, 2002, 01:48:44
Job time: 562 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:47:18 ; Search time 133.15 Seconds
(without alignments)
58.394 Million cell updates/sec

Title: US-09-289-346A-6
Perfect score: 357
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FAFAALNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	357	100.0	70	21	AAB18682	Mutant peptide der
2	340	95.2	70	21	AAB18677	Peptide fragment f
3	340	95.2	356	21	AAB18687	Amino acid sequenc
4	332	93.0	70	21	AAB18685	Mutant peptide der
5	329	92.2	70	21	AAB18688	Mutant peptide der
6	328	91.9	70	21	AAB18692	Mutant peptide der
7	326	91.3	70	21	AAB18684	Mutant peptide der
8	326	91.3	70	21	AAB18690	Mutant peptide der
9	325	91.0	70	21	AAB18678	Mutant peptide der
10	325	91.0	70	21	AAB18686	Mutant peptide der
11	324	90.8	70	21	AAB18689	Mutant peptide der
12	322	90.2	70	21	AAB18680	Mutant peptide der
13	322	90.2	70	21	AAB18691	Mutant peptide der
14	319	89.4	70	21	AAB18681	Mutant peptide der
15	318	89.1	70	21	AAB18683	Mutant peptide der
16	308	86.3	70	21	AAB18679	Mutant peptide der
17	210	58.8	359	17	AAR88870	Sardinian tomato y
18	210	58.8	359	17	AAR88871	Sardinian tomato y
19	210	58.8	359	17	AAR88872	Sardinian tomato y
20	206	57.7	353	18	AAW34338	Bean golden mosaic
21	206	57.7	353	18	AAW34332	Bean golden mosaic
22	206	57.7	353	18	AAW34333	Bean golden mosaic
23	206	57.7	353	18	AAW34334	Bean golden mosaic
24	206	57.7	353	18	AAW34335	Bean golden mosaic
25	204	57.1	353	8	AAP70407	ORF 4 gene product
26	204	57.1	361	18	AAW34336	Tomato mottle viru
27	204	57.1	361	18	AAW34324	Tomato mottle viru
28	204	57.1	361	18	AAW34325	Tomato mottle viru
29	204	57.1	361	18	AAW34326	Tomato mottle viru
30	197.5	55.3	361	8	AAP70562	Product of ORF 4 f
31	192	53.8	362	19	AAW56495	Tobacco leaf curli
32	191	53.5	357	18	AAW34329	Tomato yellow leaf
33	191	53.5	357	18	AAW34330	Tomato yellow leaf
34	191	53.5	357	18	AAW34331	Tomato yellow leaf
35	183	51.3	357	18	AAW34337	Tomato yellow leaf
36	76.5	21.4	665	22	ABB70991	Drosophila melanog
37	67	18.8	447	21	AAY99659	Human GTPase assoc
38	67	18.8	447	21	AAY87089	Human secreted pro
39	67	18.8	447	21	AAY66744	Membrane-bound pro
40	67	18.8	447	21	AAY50947	Human adult aorta
41	67	18.8	447	22	AAU29150	Human PRO polypept
42	67	18.8	447	22	AAM93267	Human polypeptide,
43	67	18.8	447	22	AAE06066	Human gene 26 enco
44	67	18.8	447	22	AAB65267	Human PRO1125 (UNQ
45	67	18.8	456	21	AAY87190	Human secreted pro

ALIGNMENTS

RESULT 1
AAB18682
ID AAB18682 standard; peptide; 70 AA.
XX
AC AAB18682;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 52
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 54
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 55
FT /note= "wild type residue replaced with Ala"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.

XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DR WPI; 2000-618851/59.
XX
PT Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT -
XX
PS Claim 53; Page 44-45; 73pp; English.
XX
CC The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
SQ Sequence 70 AA;

Query Match 100.0%; Score 357; DB 21; Length 70;
Best Local Similarity 100.0%; Pred. No. 8.9e-38;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfafaalnsnl 60

Qy 61 DRIFDKTPEP 70
||||||||||
Db 61 drifdktpep 70

RESULT 2
AAB18677
ID AAB18677 standard; peptide; 70 AA.
XX
AC AAB18677;
XX
DT 22-JAN-2001 (first entry)
XX
DE Peptide fragment from Rep protein of TGMV (amino acids 110-179).
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Tomato golden mosaic virus.
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DR WPI; 2000-618851/59.
XX
PT Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT -
XX
PS Disclosure; Page 18; 73pp; English.
XX
CC The present sequence is derived from a geminivirus replication (Rep)
CC protein, which is also known as AL1. AL1 binds double-stranded DNA,
CC catalyses cleavage and ligation of single-stranded DNA, and interacts
CC with other viral and host proteins. Mutants of the AL1 protein are used
CC to produce transgenic plants. The mutation in AL1 is present in a
CC ribosome binding region, and expression of mutant AL1 protein imparts
CC increased resistance to geminivirus infection in the plant. Mutant AL1
CC proteins are useful for producing plants having increased resistance or
CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,
CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl
CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato
CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,
CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or
CC beet curly top virus.
XX
SQ Sequence 70 AA;

Query Match 95.2%; Score 340; DB 21; Length 70;
Best Local Similarity 95.7%; Pred. No. 1.3e-35;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
|||||||||||||||||||||||||||||||||||||||||||||||||||| | |||||
Db 1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy 61 DRIFDKTPEP 70
||||||||||
Db 61 drifdktpep 70

RESULT 3
AAB18687
ID AAB18687 standard; peptide; 356 AA.
XX
AC AAB18687;
XX
DT 22-JAN-2001 (first entry)
XX
DE Amino acid sequence of a geminivirus replication protein of TGMV.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 354
FT /note= "unspecified amino acid"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DR WPI; 2000-618851/59.
XX
PT Transgenic plants with increased resistance to geminivirus infection

PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT -
XX
PS Disclosure; Page 47-48; 73pp; English.
XX
CC The present sequence represents a geminivirus replication (Rep)
CC protein, which is also known as AL1. AL1 binds double-stranded DNA,
CC catalyses cleavage and ligation of single-stranded DNA, and interacts
CC with other viral and host proteins. Mutants of the AL1 protein are used
CC to produce transgenic plants. The mutation in AL1 is present in a
CC ribosome binding region, and expression of mutant AL1 protein imparts
CC increased resistance to geminivirus infection in the plant. Mutant AL1
CC proteins are useful for producing plants having increased resistance or
CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,
CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl
CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato
CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,
CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or
CC beet curly top virus.
XX
SQ Sequence 356 AA;

Query Match 95.2%; Score 340; DB 21; Length 356;
Best Local Similarity 95.7%; Pred. No. 9.8e-35;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||| |  |||||
Db     110 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 169

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     170 drifdktpep 179
```

RESULT 4
AAB18685
ID AAB18685 standard; peptide; 70 AA.
XX
AC AAB18685;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 10
FT /note= "wild type residue replaced with Ala"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DR WPI; 2000-618851/59.
XX
PT Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT -
XX
PS Claim 53; Page 46; 73pp; English.
XX
CC The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
SQ Sequence 70 AA;

Query Match 93.0%; Score 332; DB 21; Length 70;
Best Local Similarity 94.3%; Pred. No. 1.3e-34;
Matches 66; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |||||||| ||||||||||||||||||||||||||||||||||| |  |||||
Db       1 tlvwgefqvagrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70
```

RESULT 5
AAB18688
ID AAB18688 standard; peptide; 70 AA.
XX
AC AAB18688;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 19
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 20
FT /note= "wild type residue replaced with Ala"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DR WPI; 2000-618851/59.

```
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 48; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             92.2%;  Score 329;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 3.2e-34;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |||||||||||||||||  |||||||||||||||||||||||||||| |  |||||
Db       1 tlvwgefqvdgrsarggcaasndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT   6
AAB18692
ID   AAB18692 standard; peptide; 70 AA.
XX
AC   AAB18692;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 66
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 69
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 50; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             91.9%;  Score 328;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 4.3e-34;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |||||||||||||||||||||||||||||||||||||||||||||||||| |  |||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||| || |
Db      61 drifdatpap 70


RESULT   7
AAB18684
ID   AAB18684 standard; peptide; 70 AA.
XX
AC   AAB18684;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 7
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 8
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
```

```
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;
```

```
  Query Match             91.3%;  Score 326;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 7.7e-34;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           ||||||  |||||||||||||||||||||||||||||||||||||||||| |  |||||
Db       1 tlvwgeaavdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70
```

```
RESULT   8
AAB18690
ID   AAB18690 standard; peptide; 70 AA.
XX
AC   AAB18690;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 27
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 30
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
```

```
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;
```

```
  Query Match             91.3%;  Score 326;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 7.7e-34;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           ||||||||||||||||||||||||| || |||||||||||||||||| |  |||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaaalaasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70
```

```
RESULT   9
AAB18678
ID   AAB18678 standard; peptide; 70 AA.
XX
AC   AAB18678;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 12
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 13
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 15
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
```

```
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 42-43; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             91.0%;  Score 325;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 1e-33;
  Matches   64;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |||||||||| :| ||||||||||||||||||||||||||||||||||| |  |||||
Db       1 tlvwgefqvdgaaaaggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  10
AAB18686
ID   AAB18686 standard; peptide; 70 AA.
XX
AC   AAB18686;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 24
FT                   /note= "wild type residue replaced with Leu"
FT   Misc-difference 25
FT                   /note= "wild type residue replaced with Leu"
FT   Misc-difference 26
FT                   /note= "wild type residue replaced with Leu"
XX
```

```
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 46; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             91.0%;  Score 325;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 1e-33;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |||||||||||||||||||||   |||||||||||||||||||||||| |  |||||
Db       1 tlvwgefqvdgrsarggcqtsndlllealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  11
AAB18689
ID   AAB18689 standard; peptide; 70 AA.
XX
AC   AAB18689;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 22
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 23
```

```
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 48-49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             90.8%;  Score 324;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 1.4e-33;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |||||||||||||||||||  |||||||||||||||||||||||||| |  |||||
Db       1 tlvwgefqvdgrsarggcqtsaaaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  12
AAB18680
ID   AAB18680 standard; peptide; 70 AA.
XX
AC   AAB18680;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 42
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 43
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 44
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 43-44; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             90.2%;  Score 322;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 2.5e-33;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |||||||||||||||||||||||||||||||||||||||||||   ||||||| |  |||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiiaaaipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  13
AAB18691
ID   AAB18691 standard; peptide; 70 AA.
XX
AC   AAB18691;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
```

```
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 34
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 35
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 36
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             90.2%;  Score 322;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 2.5e-33;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           ||||||||||||||||||||||||||||||||   |||||||||||||| |  |||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnassaaaalqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  14
AAB18681
ID   AAB18681 standard; peptide; 70 AA.
XX
AC   AAB18681;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 47
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 48
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 49
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 44; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             89.4%;  Score 319;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 6e-33;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||   || |  |||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipaaalfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  15
AAB18683
ID   AAB18683 standard; peptide; 70 AA.
XX
AC   AAB18683;
XX
```

```
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 59
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 61
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 62
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             89.1%;  Score 318;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 8e-33;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||| |  ||| |
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsal 60

Qy      61 DRIFDKTPEP 70
               ||||||||
Db      61 aaifdktpep 70


Search completed: October 10, 2002, 01:47:18

Job time: 527 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:01:59 ; Search time 118.38 Seconds
(without alignments)
102.295 Million cell updates/sec

Title: US-09-289-346A-7
Perfect score: 359
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSALAAIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	293	81.6	351	12	Q91R10	Q91r10 tomato seve
2	291	81.1	352	12	Q9E000	Q9e000 tomato rugo
3	274	76.3	226	12	O09727	O09727 leonurus mo
4	274	76.3	226	12	Q9WHF6	Q9whf6 tomato mild
5	272	75.8	361	12	Q67574	Q67574 bean golden
6	270	75.2	225	12	Q9QDB1	Q9qdb1 cowpea gold
7	270	75.2	314	12	Q9ELT8	Q9elt8 sweet potat
8	270	75.2	364	12	Q9QS55	Q9qs55 sweet potat
9	267	74.4	185	12	Q98693	Q98693 sida golden
10	259	72.1	149	12	P88975	P88975 macroptiliu
11	259	72.1	233	12	Q9YLA4	Q9yla4 macroptiliu
12	258	71.9	234	12	O39180	O39180 geminivirid
13	252	70.2	190	12	Q9Z089	Q9z089 tobacco lea
14	252	70.2	190	12	Q9Z084	Q9z084 tobacco lea
15	249	69.4	190	12	Q9W827	Q9w827 tobacco lea
16	249	69.4	208	12	Q9Z0C4	Q9z0c4 tobacco lea
17	249	69.4	353	12	O72692	O72692 beet curly
18	249	69.4	354	12	Q65438	Q65438 beet curly
19	249	69.4	359	12	Q91M88	Q91m88 tobacco lea
20	247	68.8	208	12	Q9Z0C0	Q9z0c0 tobacco lea
21	247	68.8	208	12	Q9Z0B8	Q9z0b8 tobacco lea
22	246	68.5	203	12	Q9Z083	Q9z083 tobacco lea
23	245	68.2	208	12	Q9Z0B6	Q9z0b6 tobacco lea
24	245	68.2	358	12	Q65418	Q65418 beet curly
25	244	68.0	363	12	O73577	O73577 cotton leaf
26	239	66.6	359	12	Q91B86	Q91b86 ageratum ye
27	239	66.6	360	12	Q9DXI0	Q9dxi0 ageratum ye
28	238	66.3	190	12	Q9Z086	Q9z086 tobacco lea
29	238	66.3	348	12	Q911W5	Q911w5 macroptiliu
30	238	66.3	362	12	O56816	O56816 chayote mos
31	237	66.0	208	12	Q9Z0A0	Q9z0a0 tobacco lea
32	237	66.0	362	12	Q9IN48	Q9in48 okra enatio
33	237	66.0	363	12	O72705	O72705 cotton leaf
34	237	66.0	363	12	O72719	O72719 cotton leaf
35	236	65.7	208	12	Q9Z0C6	Q9z0c6 tobacco lea
36	235	65.5	349	12	Q88888	Q88888 tomato pseu
37	234	65.2	359	12	Q9YZV4	Q9yzv4 tomato yell
38	234	65.2	359	12	Q9YZV2	Q9yzv2 tomato yell
39	234	65.2	359	12	Q9YUX7	Q9yux7 tomato yell
40	234	65.2	359	12	Q9YL27	Q9yl27 tomato yell
41	233	64.9	359	12	Q88942	Q88942 tomato yell
42	232	64.6	190	12	Q9Z0A7	Q9z0a7 tobacco lea
43	231	64.3	307	12	Q9IET7	Q9iet7 cotton leaf
44	231	64.3	361	12	O72723	O72723 cotton leaf
45	230	64.1	231	12	Q96620	Q96620 african tom

ALIGNMENTS

RESULT 1
Q91R10
ID Q91R10 PRELIMINARY; PRT; 351 AA.
AC Q91R10;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE REP PROTEIN.
GN AC1.
OS Tomato severe rugose virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=158463;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MINAS GERAIS;
RA Rezende W.L., Goulart L.R., Parreira K.S., Figueiredo J.E.F.;
RT "The full-length DNA-A nucleotide sequence of a novel tomato-infecting
RT begomovirus, Tomato severe rugose virus, in Brazil.";
RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY029750; AAK50357.1; -.
SQ SEQUENCE 351 AA; 40122 MW; 87F937A4F873B6CF CRC64;

Query Match 81.6%; Score 293; DB 12; Length 351;
Best Local Similarity 78.6%; Pred. No. 2.9e-25;
Matches 55; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: ||||:|||||||||||:|||||||||| ||: ||||||||:|||:|||||||||| |
Db     111 TIEWGEFQIDGRSARGGCQTANDAAAEALNAPSKDVALQIIREKLPEKFLFQFHNLNSNL 170

Qy      61 AAIFDKTPEP 70
             || : |||
Db     171 DRIFARAPEP 180
```

RESULT 2
Q9E000
ID Q9E000 PRELIMINARY; PRT; 352 AA.

```
AC   Q9E000;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN.
GN   AC1.
OS   Tomato rugose mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=134599;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Fernandes J.J., Fontes E.P.B., Brommonschenkel S.H., Carvalho M.G.,
RA   Zambolim E.M., Zerbini F.M.;
RT   "Molecular Cloning and Characterization of Tomato rugose mosaic virus
RT   (TRMV), a Begomovirus Isolated from Tomatoes at Triangulo Mineiro,
RT   Minas Gerais, Brazil.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF291705; AAG15546.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   352 AA;  40012 MW;  47CD55838E24D613 CRC64;


  Query Match             81.1%;  Score 291;  DB 12;  Length 352;
  Best Local Similarity   78.6%;  Pred. No. 4.9e-25;
  Matches   55;  Conservative    7;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: |||||:||||||||||:|||||||||| ||: ||||||||:|||:||||||||| |
Db     111 TIEWGEFQIDGRSARGGCQTANDAAAEALNAPSKDIALQIIREKMPEKFLFQFHNLNSNL 170

Qy      61 AAIFDKTPEP 70
             || : |||
Db     171 DRIFARAPEP 180


RESULT   3
O09727
ID   O09727      PRELIMINARY;      PRT;   226 AA.
AC   O09727;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   REP.
OS   Leonurus mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=58177;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LEMV- BRAZIL 1;
RA   Faria J.C., Maxwell D.P.;
RT   "Variability in geminivirus associated with Phaseolus vulgaris in
RT   Brazil.";
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U92532; AAB51157.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     226    226
SQ   SEQUENCE   226 AA;  25617 MW;  73CDB6E766083FC5 CRC64;


  Query Match             76.3%;  Score 274;  DB 12;  Length 226;
  Best Local Similarity   77.1%;  Pred. No. 2.5e-23;
  Matches   54;  Conservative    4;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: ||||||||||:||| || |||||||||| |  ||||||:||:|||||||||||||| |
Db     111 TVEWGEFQVDGRSSRGGQQTVNDAAAEALNAPDKRTALQIIKEKLPEKYLFQFHNLNSNL 170

Qy      61 AAIFDKTPEP 70
             || | |||
Db     171 DRIFAKAPEP 180


RESULT   4
Q9WHF6
ID   Q9WHF6      PRELIMINARY;      PRT;   226 AA.
AC   Q9WHF6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   tomato mild mottle geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX   NCBI_TaxID=92943;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HN96-H5KW;
RA   Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA   Maxwell D.P.;
RT   "Molecular characterization and DNA-based detection methods for
RT   vegetable-infecting geminiviruses in Central America.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF131071; AAD33471.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     226    226
SQ   SEQUENCE   226 AA;  25941 MW;  2EA4116712871A23 CRC64;


  Query Match             76.3%;  Score 274;  DB 12;  Length 226;
  Best Local Similarity   72.9%;  Pred. No. 2.5e-23;
  Matches   51;  Conservative   12;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: |||||:|||||||| ||:||||||||||||||||::||:||:|:::|||:|||:| |
Db     111 TIEWGEFQIDGRSARGGQQTANDAAAEALNASSKEEAMRIIKEKLPKEFLFQYHNLSSNL 170

Qy      61 AAIFDKTPEP 70
             || | |||
Db     171 DRIFAKAPEP 180


RESULT   5
Q67574
ID   Q67574      PRELIMINARY;      PRT;   361 AA.
AC   Q67574;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE  REPLICATIVE PROTEIN.
GN   AL1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,
RA   Maxwell D.P., Russell D.R.;
RT   "Cloning of the complete DNA genomes of four bean-infecting
RT   geminiviruses and determining their infectivity by electric discharge
RT   particle acceleration.";
RL   Phytopathology 81:980-985(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,
RA   Morales F.J., Maxwell D.P.;
RT   "Differentiation of bean-infecting geminiviruses by nucleic acid
```

```
RT   hybridization probes and aspects of bean golden mosaic in Brazil.";
RL   Plant Dis. 75:336-342(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RT   "Genetic diversity in geminiviruses causing bean golden mosaic
RT   disease: The nucleotide sequence of the infectious cloned DNA
RT   components of a Brazilian isolate of bean golden mosaic geminivirus.";
RL   Submitted (MAY-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; M88686; AAA46312.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   361 AA;  41041 MW;  0094C7ACAF06B788 CRC64;


  Query Match           75.8%;  Score 272;  DB 12;  Length 361;
  Best Local Similarity  76.1%;  Pred. No. 7.3e-23;
  Matches   51;  Conservative    9;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREK1PEKYLFQFHNLNSALAAI 63
           || |||||||||| ||:||||:||||||||||:|||:||:|||:|||:| |  |
Db     113 WGHFQVDGRSARGGQQTANDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
           | | |:|
Db     173 FTKAPDP 179


RESULT   6
Q9QDB1
ID   Q9QDB1      PRELIMINARY;      PRT;   225 AA.
AC   Q9QDB1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   cowpea golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=69263;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CGMV-BR;
RA   Faria J.C.;
RT   "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT   Brazil.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF188708; AAF06318.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     225    225
SQ   SEQUENCE   225 AA;  25766 MW;  1089CB6BD8D15B5D CRC64;


  Query Match           75.2%;  Score 270;  DB 12;  Length 225;
  Best Local Similarity  76.1%;  Pred. No. 7.2e-23;
  Matches   51;  Conservative    8;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALAAI 63
           || ||:|||||||| || ||||:||||||||||:|||:||:|||:|||:| |  |
Db     113 WGHFQIDGRSARGGQQTINDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
           | | |||
Db     173 FKKPPEP 179


RESULT   7
Q9ELT8
ID   Q9ELT8      PRELIMINARY;      PRT;   314 AA.
AC   Q9ELT8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION ASSOCIATION PROTEIN.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF288227; AAG01006.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   314 AA;  35153 MW;  6B6220613046943F CRC64;


  Query Match           75.2%;  Score 270;  DB 12;  Length 314;
  Best Local Similarity  79.4%;  Pred. No. 1e-22;
  Matches   54;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: |||||||||||||| ||:|||||||||| ||| ||||||||:|||||||||||| | |
Db     110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy      61 AAIFDKTP 68
             ||   |
Db     170 DRIFSPPP 177


RESULT   8
Q9QS55
ID   Q9QS55      PRELIMINARY;      PRT;   364 AA.
AC   Q9QS55;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION INITIATION PROTEIN AC1.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A.;
RT   "Cloning of a DNA-A-like genomic component of sweet potato leaf curl
RT   virus : nucleotide sequence and phylogenetic relationships.";
RL   Online Publication.
DR   EMBL; AF104036; AAD47173.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   364 AA;  40680 MW;  5F79752431A09D6E CRC64;
```

```
  Query Match             75.2%;  Score 270;  DB 12;  Length 364;
  Best Local Similarity   79.4%;  Pred. No. 1.2e-22;
  Matches   54;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: |||||||||||||| ||:|||||||||| ||| ||||||||:|||||||||| | |
Db     110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy      61 AAIFDKTP 68
             ||   |
Db     170 DRIFSPPP 177


RESULT   9
Q98693
ID   Q98693      PRELIMINARY;      PRT;   185 AA.
AC   Q98693;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   AC1.
OS   sida golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51034;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA;
RA   Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RT   "Genetic Diversity among geminiviruses associated with the weed
RT   species Sida spp, Macroptilium lathyroides, and Wissadula amplissima
RT   from Jamaica.";
RL   Plant Dis. 81:1251-1258(1997).
DR   EMBL; U67926; AAB97865.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     185    185
SQ   SEQUENCE   185 AA;  20975 MW;  3913850A025A5EE1 CRC64;


  Query Match             74.4%;  Score 267;  DB 12;  Length 185;
  Best Local Similarity   70.0%;  Pred. No. 1.3e-22;
  Matches   49;  Conservative   12;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: || ||:|||||||| ||:|||||||||: :||:||:|||||:||:||||:|||:| :
Db      89 TIEWGVFQIDGRSARGGQQTANDAAAEALNSGTKEDALKIIREKLPERYLFQYHNLSSNI 148

Qy      61 AAIFDKTPEP 70
             || | |||
Db     149 DRIFSKPPEP 158


RESULT  10
P88975
ID   P88975      PRELIMINARY;      PRT;   149 AA.
AC   P88975;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   AC1.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICAN;
RA   Roye M.E.;
RL   Thesis (1996), Biochemistry, University of the West Indies, Jamaica.
DR   EMBL; U75278; AAB36919.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     149    149
SQ   SEQUENCE   149 AA;  16785 MW;  E4CF5EED4C9CD508 CRC64;


  Query Match             72.1%;  Score 259;  DB 12;  Length 149;
  Best Local Similarity   68.6%;  Pred. No. 8e-22;
  Matches   48;  Conservative   12;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: || ||:|||||||| ||||||||||||: :|| |::|::||:|||:|||:|||:| |
Db      52 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 111

Qy      61 AAIFDKTPEP 70
             || | |||    .
Db     112 DRIFMKDPEP 121


RESULT  11
Q9YLA4
ID   Q9YLA4      PRELIMINARY;      PRT;   233 AA.
AC   Q9YLA4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E.;
RT   "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
RT   from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E., McLaughlin W.A., Maxwell D.P.;
RT   "Molecular characterization of two distinct geminiviruses infecting M.
RT   lathyroides from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF098940; AAD17850.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     233    233
SQ   SEQUENCE   233 AA;  26356 MW;  AA490AF4D2166A02 CRC64;


  Query Match             72.1%;  Score 259;  DB 12;  Length 233;
  Best Local Similarity   68.6%;  Pred. No. 1.3e-21;
  Matches   48;  Conservative   12;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: || ||:|||||||| ||||||||||||: :|| |::|::||:|||:|||:|||:| |
Db     110 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 169

Qy      61 AAIFDKTPEP 70
             || | |||
Db     170 DRIFMKDPEP 179


RESULT  12
O39180
```

```
ID   O39180      PRELIMINARY;      PRT;   234 AA.
AC   O39180;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   AC1 (FRAGMENT).
OS   Geminiviridae.
OC   Viruses; ssDNA viruses.
OX   NCBI_TaxID=10811;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOMATO STRAIN;
RA   Guzman P., Arredondo C.R., Emmatty D., Portillo R.J., Gilbertson R.L.;
RT   "Partial Characterization of Two Whitefly-Transmitted Geminiviruses
RT   Infecting Tomatoes in Venezuela.";
RL   Plant Dis. 81:312-312(1997).
DR   EMBL; AF026553; AAB82605.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     234    234
SQ   SEQUENCE   234 AA;  26486 MW;  9ED8F0697105CD19 CRC64;


  Query Match             71.9%;  Score 258;  DB 12;  Length 234;
  Best Local Similarity   68.6%;  Pred. No. 1.7e-21;
  Matches   48;  Conservative   12;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: || ||:|||||||| ||:|||||||||: :||||::||:||:|||:|||:|||:| |
Db     110 TIEWGLFQIDGRSARGGQQTANDAAAEALNSGTKEEAMKIIKEKLPEKFLFQYHNLSSDL 169

Qy      61 AAIFDKTPEP 70
             || | | |
Db     170 DRIFMKAPNP 179


RESULT  13
Q9Z089
ID   Q9Z089      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z089;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA3-1, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001315; BAA34033.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21432 MW;  AAC093D1D1610FAD CRC64;


  Query Match             70.2%;  Score 252;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 6.6e-21;
  Matches   50;  Conservative    9;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           || || ||:|||||||||| :||| |||||||||| ||| |||||:|: ::||:||||| |
Db      85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy      61 ----------------AAIFDKTPE 69
                            |: ||: ||
Db     145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  14
Q9Z084
ID   Q9Z084      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z084;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA5-2, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA5;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001318; BAA34039.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  AAC1C2943E3F01AD CRC64;


  Query Match             70.2%;  Score 252;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 6.6e-21;
  Matches   50;  Conservative    9;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           || || ||:|||||||||| :||| |||||||||| ||| |||||:|: ::||:||||| |
Db      85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy      61 ----------------AAIFDKTPE 69
                            |: ||: ||
Db     145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  15
Q9W827
ID   Q9W827      PRELIMINARY;      PRT;   190 AA.
AC   Q9W827;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE GORA1-1, PARTIAL AND
DE   COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GORA1;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001303; BAA34010.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
```

```
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  93C3742A8EBDB7EB CRC64;


  Query Match             69.4%;  Score 249;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 1.5e-20;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           || || |||||||||||| :||| |||||||:|| ||| |||||:|: ::||:||||| |
Db      85 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 144

Qy      61 AAI----------------FDKTPE 69
             |                ||: ||
Db     145 DRIFAPPLEVFVCPFSNSSFDQVPE 169


Search completed: October 10, 2002, 02:02:00
Job time: 1093 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:09:52 ; Search time 34.62 Seconds
(without alignments)
78.289 Million cell updates/sec

Title: US-09-289-346A-7
Perfect score: 359
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSALAAIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	342	95.3	352	1	VAL1_TGMV	P03567 tomato gold
2	242	67.4	361	1	VAL1_PYMVV	P27258 potato yell
3	229	63.8	358	1	VAL1_CLVK	P14982 cassava lat
4	229	63.8	358	1	VAL1_CLVN	P14972 cassava lat
5	222	61.8	362	1	VAL1_TYLCA	P36279 tomato yell
6	216	60.2	349	1	VAL1_PHUV	Q06923 pepper huas
7	215	59.9	359	1	VAL1_TYLCU	P38609 tomato yell
8	212	59.1	358	1	VAL1_BCTV	P14991 beet curly
9	211	58.8	353	1	VAL1_BGMV	P05175 bean golden
10	211	58.8	359	1	VAL1_TYLCM	P27260 tomato yell
11	210	58.5	355	1	VAL1_ABMVW	P21947 abutilon mo
12	207	57.7	361	1	VAL1_TMOV	Q06657 tomato mott
13	193	53.8	357	1	VAL1_TYLCV	P27259 tomato yell
14	107	29.8	347	1	VAL1_SLCV	P29048 squash leaf
15	68	18.9	447	1	TBL2_HUMAN	Q9y4p3 homo sapien
16	63.5	17.7	1610	1	CCAD_MESAU	Q99244 mesocricetu
17	63.5	17.7	2161	1	CCAD_HUMAN	Q01668 homo sapien
18	63.5	17.7	2203	1	CCAD_RAT	P27732 rattus norv
19	61	17.0	1483	1	CYP1_YEAST	P12351 saccharomyc
20	59.5	16.6	224	1	SP2R_BACSU	P39151 bacillus su
21	59.5	16.6	4466	1	DYHC_TRIGR	P23098 tripneustes
22	59	16.4	387	1	VE2_HPV41	P27552 human papil
23	59	16.4	387	1	Y4PF_RHISN	P55615 rhizobium s
24	58.5	16.3	136	1	Y452_CAEEL	O62250 caenorhabdi
25	58.5	16.3	299	1	Y175_HELPJ	Q9zmq7 helicobacte
26	57.5	16.0	367	1	LHX4_MOUSE	P53776 mus musculu
27	57.5	16.0	1852	1	CCAS_CYPCA	P22316 cyprinus ca
28	57	15.9	247	1	YCP4_YEAST	P25349 saccharomyc
29	57	15.9	256	1	YKJ9_YEAST	P34247 saccharomyc
30	56.5	15.7	511	1	HUTH_VIBCH	Q9ksq4 vibrio chol
31	56.5	15.7	2190	1	CCAD_CHICK	O73700 gallus gall
32	56	15.6	98	1	VATF_METJA	Q57671 methanococc
33	56	15.6	295	1	VAL1_CSMV	P18921 chloris str
34	56	15.6	295	1	VAL1_TYDVA	P31617 tobacco yel
35	56	15.6	492	1	MOT3_MOUSE	O35308 mus musculu
36	56	15.6	492	1	MOT3_RAT	O70461 rattus norv
37	56	15.6	836	1	GCSR_HUMAN	Q99062 homo sapien
38	55.5	15.5	266	1	ETFB_MYCLE	O33095 mycobacteri
39	55.5	15.5	266	1	ETFB_MYCTU	O53276 mycobacteri
40	55.5	15.5	428	1	GBA1_CANAL	P28868 candida alb
41	55.5	15.5	887	1	MTP_BOVIN	P55156 bos taurus
42	55.5	15.5	1066	1	Z295_HUMAN	Q9ulj3 homo sapien
43	55.5	15.5	4466	1	DYHC_ANTCR	P39057 anthocidari
44	55	15.3	354	1	CARA_METJA	Q58425 methanococc
45	55	15.3	481	1	KNIR_DROVI	Q24753 drosophila

ALIGNMENTS

RESULT 1
VAL1_TGMV
ID VAL1_TGMV STANDARD; PRT; 352 AA.
AC P03567;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein.
GN AC1.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10831;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;
RT "Complete nucleotide sequence of the infectious cloned DNA components
RT of tomato golden mosaic virus: potential coding regions and regulatory
RT sequences.";
RL EMBO J. 3:2197-2205(1984).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; K02029; -; NOT_ANNOTATED_CDS.
DR PIR; A04170; QQCVL1.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR ProDom; PD000736; Gemini_AL1; 1.
KW ATP-binding.
FT NP_BIND 223 230 ATP (POTENTIAL).
SQ SEQUENCE 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;

Query Match 95.3%; Score 342; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 1.7e-31;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 AAIFDKTPEP 70
           ||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
VAL1_PYMVV
ID VAL1_PYMVV STANDARD; PRT; 361 AA.

```
AC   P27258;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 protein.
OS   Potato yellow mosaic virus (isolate Venezuela).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10828;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91311403; PubMed=1856690;
RA   Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT   "The nucleotide sequence of the infectious cloned DNA components of
RT   potato yellow mosaic virus.";
RL   J. Gen. Virol. 72:1515-1520(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; D00940; BAA00782.1; -.
DR   PIR; JU0364; QQCVPT.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   361 AA;  40850 MW;  5627A33BF1264383 CRC64;


  Query Match             67.4%;  Score 242;  DB 1;  Length 361;
  Best Local Similarity   66.7%;  Pred. No. 3.8e-20;
  Matches   46;  Conservative   11;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: || ||:|||||||| || |||||||||: :|| |::||:||:|||:|||:|||:  |
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 AAIFDKTPE 69
             || | ||
Db     170 DRIFMKAPE 178


RESULT   3
VAL1_CLVK
ID   VAL1_CLVK      STANDARD;      PRT;   358 AA.
AC   P14982;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (40.4 kDa protein).
GN   AC1.
OS   Cassava latent virus (strain West Kenyan 844).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10818;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Gay M.R.;
RT   "Nucleotide sequence of cassava latent virus DNA.";
RL   Nature 301:260-262(1983).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; J02057; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40346 MW;  ED173E753EE92D69 CRC64;


  Query Match             63.8%;  Score 229;  DB 1;  Length 358;
  Best Local Similarity   60.0%;  Pred. No. 1.1e-18;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: ||:||:|||||||| |::||| |:|||: || ||| :||| :|: :: |||||| |
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 AAIFDKTPEP 70
             || : | |
Db     169 DRIFQEPPAP 178


RESULT   4
VAL1_CLVN
ID   VAL1_CLVN      STANDARD;      PRT;   358 AA.
AC   P14972;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (40.4 kDa protein).
GN   AC1.
OS   Cassava latent virus (strain Nigerian).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10819;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90174930; PubMed=2308831;
RA   Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT   "Nucleotide sequence of the infectious cloned DNA components of
RT   African cassava mosaic virus (Nigerian strain).";
RL   Nucleic Acids Res. 18:197-198(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X17095; CAA34953.1; -.
DR   PIR; S07594; S07594.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40435 MW;  1DB16BB0CB2D5E2C CRC64;


  Query Match             63.8%;  Score 229;  DB 1;  Length 358;
  Best Local Similarity   60.0%;  Pred. No. 1.1e-18;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: ||:||:|||||||| |::||| |:|||: || ||| :||| :|: :: |||||| |
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168
```

```
Qy      61 AAIFDKTPEP 70
           || : | |
Db     169 DRIFQEPPAP 178


RESULT   5
VAL1_TYLCA
ID   VAL1_TYLCA     STANDARD;      PRT;   362 AA.
AC   P36279;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=36447;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93139778; PubMed=8423446;
RA   Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;
RT   "Nucleotide sequence and genome organization of tomato leaf curl
RT   geminivirus.";
RL   J. Gen. Virol. 74:147-151(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
DR   PIR; JQ1887; JQ1887.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   362 AA;  41197 MW;  343E7184B4704098 CRC64;


  Query Match             61.8%;  Score 222;  DB 1;  Length 362;
  Best Local Similarity   52.9%;  Pred. No. 7.1e-18;
  Matches   45;  Conservative    9;  Mismatches   15;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
             || |||||:|||||||| |::||| |:|||  || ||| ::||  |: |: ||||||| |
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 AAI----------------FDKTPE 69
             |                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   6
VAL1_PHUV
ID   VAL1_PHUV      STANDARD;      PRT;   349 AA.
AC   Q06923;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AL1.
OS   Pepper huasteco virus (PHV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=28349;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94015007; PubMed=8409944;
RA   Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA   Rivera-Bustamante R.F.;
RT   "Complete nucleotide sequence of pepper huasteco virus: analysis and
RT   comparison with bipartite geminiviruses.";
RL   J. Gen. Virol. 74:2225-2231(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70418; CAA49856.1; -.
DR   PIR; S31875; S31875.
DR   PIR; JQ2300; JQ2300.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (BY SIMILARITY).
SQ   SEQUENCE   349 AA;  39722 MW;  D5F4E76CD56370F4 CRC64;


  Query Match             60.2%;  Score 216;  DB 1;  Length 349;
  Best Local Similarity   58.6%;  Pred. No. 3.3e-17;
  Matches   41;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
             |: |||||:|||||||| |::||  |:|||::| |||||||:|: |: :  ||||: |
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 AAIFDKTPEP 70
             ||   |||
Db     170 NRIFQTPPEP 179


RESULT   7
VAL1_TYLCU
ID   VAL1_TYLCU     STANDARD;      PRT;   359 AA.
AC   P38609;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=37139;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94256836; PubMed=8198442;
RA   Noris E., Hidalgo E., Accotto G., Moriones E.;
RT   "High similarity among the tomato yellow leaf curl virus isolates
RT   from the west Mediterranean basin: the nucleotide sequence of an
RT   infectious clone from Spain.";
RL   Arch. Virol. 135:165-170(1994).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z25751; CAA81026.1; -.
DR   PIR; S39211; S39211.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  41065 MW;  2D170A51EF80A3EC CRC64;


  Query Match             59.9%;  Score 215;  DB 1;  Length 359;
```

```
  Best Local Similarity   58.0%;  Pred. No. 4.4e-17;
  Matches   40;  Conservative    9;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALA 61
           | || ||:|||||||| ||:||| |:|:|| || ||| :|:|  |  |:  |||:|| |
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 AIFDKTPEP 70
             :|   | |
Db     171 RVFQVPPAP 179


RESULT   8
VAL1_BCTV
ID   VAL1_BCTV      STANDARD;      PRT;   358 AA.
AC   P14991;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 protein (40.8 kDa protein).
OS   Beet curly top virus (BCTV).
OC   Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
OX   NCBI_TaxID=10840;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT   "The nucleotide sequence of an infectious clone of the geminivirus
RT   beet curly top virus.";
RL   EMBO J. 5:1761-1767(1986).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40889 MW;  39A45FE3C0B9C333 CRC64;


  Query Match             59.1%;  Score 212;  DB 1;  Length 358;
  Best Local Similarity   55.7%;  Pred. No. 9.5e-17;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: |||||:|||||||| ||:||: |:|||||:| ::||||::|: |: |  | ||| :
Db     110 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 169

Qy      61 AAIFDKTPEP 70
             || : |:|
Db     170 QKIFQRPPDP 179


RESULT   9
VAL1_BGMV
ID   VAL1_BGMV      STANDARD;      PRT;   353 AA.
AC   P05175;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   AL1 protein (40.2 kDa protein).
GN   AC1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Howarth A.J., Caton J., Bossert M., Goodman R.M.;
RT   "Nucleotide sequence of bean golden mosaic virus and a model for gene
RT   regulation in geminiviruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M10070; AAA46318.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   353 AA;  40190 MW;  80FA779DF6029A34 CRC64;


  Query Match             58.8%;  Score 211;  DB 1;  Length 353;
  Best Local Similarity   58.6%;  Pred. No. 1.2e-16;
  Matches   41;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: ||:|||||||||||| |::||: |:|||| | | || |::|: |: |: | ||: | |
Db     110 TIEWGQFQVDGRSARGGQQSANDSYAKALNADSIESALTILKEEQPKDYVLQNHNIRSNL 169

Qy      61 AAIFDKTPEP 70
             || | |||
Db     170 ERIFFKVPEP 179


RESULT  10
VAL1_TYLCM
ID   VAL1_TYLCM     STANDARD;      PRT;   359 AA.
AC   P27260;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10833;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92107660; PubMed=1840676;
RA   Kheyr-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,
RA   Gronenborn B.;
RT   "Tomato yellow leaf curl virus from Sardinia is a
RT   whitefly-transmitted monopartite geminivirus.";
RL   Nucleic Acids Res. 19:6763-6769(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X61153; CAA43466.1; -.
DR   PIR; S22593; S22593.
DR   InterPro; IPR001191; Gemini_AL1.
```

```
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  40733 MW;  9717B4A07C93EFA7 CRC64;


  Query Match             58.8%;  Score 211;  DB 1;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 1.2e-16;
  Matches   39;  Conservative   10;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALA 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|| |
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 AIFDKTPEP 70
             :|   | |
Db     171 KVFQVPPAP 179


RESULT  11
VAL1_ABMVW
ID   VAL1_ABMVW     STANDARD;      PRT;   355 AA.
AC   P21947;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AC1.
OS   Abutilon mosaic virus (isolate West India).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10816;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91020984; PubMed=2219703;
RA   Frischmuth T., Zimmat G., Jeske H.;
RT   "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT   as well as eukaryotic features.";
RL   Virology 178:461-468(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR   PIR; A36214; QQCVW1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   355 AA;  40257 MW;  16A2CA8A63251E95 CRC64;


  Query Match             58.5%;  Score 210;  DB 1;  Length 355;
  Best Local Similarity   57.1%;  Pred. No. 1.6e-16;
  Matches   40;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |  |||||:|||||||| ||:||: |:|||    : || |::|: |: |: | ||: | |
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 AAIFDKTPEP 70
             || | |||
Db     170 ERIFAKAPEP 179


RESULT  12
VAL1_TMOV
ID   VAL1_TMOV      STANDARD;      PRT;   361 AA.
AC   Q06657;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AL1.
OS   Tomato mottle virus (isolate Florida) (TMoV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=36449;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93107858; PubMed=1469361;
RA   Abouzid A.M., Polston J.E., Hiebert E.;
RT   "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT   isolated from tomatoes in Florida.";
RL   J. Gen. Virol. 73:3225-3229(1992).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L14460; AAC32414.1; -.
DR   PIR; JQ1870; JQ1870.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (BY SIMILARITY).
SQ   SEQUENCE   361 AA;  40516 MW;  8138B65CEEAC6950 CRC64;


  Query Match             57.7%;  Score 207;  DB 1;  Length 361;
  Best Local Similarity   54.3%;  Pred. No. 3.5e-16;
  Matches   38;  Conservative   15;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: ||:||:|||||||| |::||: |:|||| | : || ::||: |: :: | ||: | |
Db     110 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 169

Qy      61 AAIFDKTPEP 70
             || | |||
Db     170 ERIFAKAPEP 179


RESULT  13
VAL1_TYLCV
ID   VAL1_TYLCV     STANDARD;      PRT;   357 AA.
AC   P27259;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10832;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92024070; PubMed=1926771;
RA   Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;
RT   "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RT   with a single genomic component.";
RL   Virology 185:151-161(1991).
```

```
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15656; CAA33688.1; -.
DR   PIR; D40779; QQCVC1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     219    226       ATP (POTENTIAL).
SQ   SEQUENCE   357 AA;  40678 MW;  939AB68E1AB3B2A7 CRC64;


  Query Match             53.8%;  Score 193;  DB 1;  Length 357;
  Best Local Similarity   62.3%;  Pred. No. 1.4e-14;
  Matches   38;  Conservative   10;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALAAI 63
           :|  |:|||||||| |::||| |||||: || ||| |::|| |: |: |||||:| |  |
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171


RESULT  14
VAL1_SLCV
ID   VAL1_SLCV      STANDARD;      PRT;   347 AA.
AC   P29048;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   AL1 protein.
OS   Squash leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10829;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91082449; PubMed=1984668;
RA   Lazarowitz S.G., Lazdins I.B.;
RT   "Infectivity and complete nucleotide sequence of the cloned genomic
RT   components of a bipartite squash leaf curl geminivirus with a broad
RT   host range phenotype.";
RL   Virology 180:58-69(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M38183; AAC32410.1; ALT_INIT.
DR   PIR; C36785; QQCVS1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     218    225       ATP (POTENTIAL).
SQ   SEQUENCE   347 AA;  39110 MW;  AFDAEBDDE122110E CRC64;


  Query Match             29.8%;  Score 107;  DB 1;  Length 347;
  Best Local Similarity   34.8%;  Pred. No. 7.6e-05;
  Matches   23;  Conservative   13;  Mismatches   26;  Indels    4;  Gaps    1;

Qy       5 GEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALAAIF 64
           |:::| |    |     :|    |:|| |  ||| ||:   |: ::  :||| : :  :|
Db     116 GQYKVSG----GSKSNKDDVYHNAVNAGSAGEALDIIKAGDPKTFIVNYHNLLANVERLF 171

Qy      65 DKTPEP 70
             | |||
Db     172 QKPPEP 177


RESULT  15
TBL2_HUMAN
ID   TBL2_HUMAN     STANDARD;      PRT;   447 AA.
AC   Q9Y4P3; Q9UQE2;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Transducin beta-like 2 protein (WS beta-transducin repeats protein)
DE   (WS-betaTRP).
GN   TBL2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20044626; PubMed=10575226;
RA   Perez Jurado L.A., Wang Y.-K., Francke U., Cruces J.;
RT   "TBL2, a novel transducin family member in the WBS deletion:
RT   characterization of the complete sequence, genomic structure,
RT   transcriptional variants and the mouse ortholog.";
RL   Cytogenet. Cell Genet. 86:277-284(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=21154917; PubMed=11230166;
RA   Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA   Ansorge W., Boecher M., Bloecker H., Bauersachs S., Blum H.,
RA   Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
RA   Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA   Wambutt R., Korn B., Klein M., Poustka A.;
RT   "Towards a catalog of human genes and proteins: sequencing and
RT   analysis of 500 novel complete protein coding human cDNAs.";
RL   Genome Res. 11:422-435(2001).
RN   [3]
RP   SEQUENCE OF 14-447 FROM N.A.
RX   MEDLINE=99075645; PubMed=9860302;
RA   Meng X., Lu X., Li Z., Green E.D., Massa H., Trask B.J., Morris C.A.,
RA   Keating M.T.;
RT   "Complete physical map of the common deletion region in Williams
RT   syndrome and identification and characterization of three novel
RT   genes.";
RL   Hum. Genet. 103:590-599(1998).
CC   -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF097484; AAF06823.1; -.
DR   EMBL; AF097485; AAF06824.1; -.
DR   EMBL; AL080162; CAB45751.1; -.
DR   EMBL; AF056183; AAD28083.1; ALT_INIT.
DR   MIM; 605842; -.
DR   InterPro; IPR001680; WD40.
```

```
DR   Pfam; PF00400; WD40; 5.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   SMART; SM00320; WD40; 5.
DR   PROSITE; PS00678; WD_REPEATS_1; 1.
DR   PROSITE; PS50082; WD_REPEATS_2; 3.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Repeat; WD repeat.
FT   REPEAT       88    127       WD 1.
FT   REPEAT      134    174       WD 2.
FT   REPEAT      186    226       WD 3.
FT   REPEAT      228    267       WD 4.
FT   REPEAT      277    316       WD 5.
FT   REPEAT      329    367       WD 6.
FT   REPEAT      371    409       WD 7.
SQ   SEQUENCE   447 AA;  49797 MW;  B260087E1A71D3F9 CRC64;


  Query Match             18.9%;  Score 68;  DB 1;  Length 447;
  Best Local Similarity   32.7%;  Pred. No. 2.7;
  Matches   16;  Conservative   11;  Mismatches   22;  Indels    0;  Gaps    0;

Qy      12 RSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           || |  || :|    :  : | |::  | ||:: |::: |    | :||
Db      38 RSGRPACQKANGFPPDKSSGSKKQKQYQRIRKEKPQQHNFTHRLLAAAL 86


Search completed: October 10, 2002, 02:09:52
Job time: 1350 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:48:44 ; Search time 69.63 Seconds
(without alignments)
96.600 Million cell updates/sec

Title: US-09-289-346A-7
Perfect score: 359
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSALAAIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	342	95.3	352	1	QQCVL1	AL1 protein - toma
2	242	67.4	361	1	QQCVPT	AL1 protein - pota
3	229	63.8	358	2	S07594	hypothetical prote
4	222	61.8	362	1	JQ1887	AL1 protein - toma
5	216	60.2	349	2	JQ2300	replicase - pepper
6	216	60.2	349	2	S31875	AL1 protein - pepp
7	215	59.9	359	2	S39211	gene C1 protein -
8	212	59.1	385	2	S28360	AL1 protein - beet
9	211	58.8	359	2	S22593	hypothetical prote
10	210	58.5	355	1	QQCVW1	AV1 protein - abut
11	207	57.7	358	1	JQ1870	AL1 protein - toma
12	205	57.1	351	2	JQ2327	AL1 protein - Indi
13	205	57.1	359	2	S39235	gene C1 protein -
14	195	54.3	360	2	S59885	replication-associ
15	193	53.8	357	1	QQCVC1	AL1 protein - toma
16	120	33.4	131	2	S45059	AC1 protein (clone
17	107	29.8	347	1	QQCVS1	AL1 protein - squa
18	70	19.5	587	2	JC1419	Fc gamma (IgG) rec
19	68	18.9	447	2	T12544	hypothetical prote
20	64.5	18.0	491	2	AE1836	leucine aminopepti
21	64	17.8	316	2	C82085	conserved hypothet
22	64	17.8	340	2	B97011	probably O-sialogl
23	63.5	17.7	1229	2	T48959	kinesin-like prote
24	63.5	17.7	1610	2	A46227	voltage-dependent
25	63.5	17.7	1646	2	JH0422	voltage-dependent
26	63.5	17.7	2161	2	JH0564	calcium channel al
27	63.5	17.7	2181	2	A38198	calcium channel al
28	63.5	17.7	2203	2	T42742	voltage-dependent
29	62	17.3	154	2	AD3475	nitrogen regulator
30	62	17.3	245	2	S76632	hypothetical prote
31	62	17.3	714	2	C95382	probable ferrichro
32	61	17.0	1502	1	RGBYH1	CYC1/CYP3 transcri
33	60.5	16.9	276	2	AC2341	hypothetical prote
34	60.5	16.9	840	2	T36175	probable large ATP
35	59.5	16.6	208	2	E97702	hypothetical prote
36	59.5	16.6	224	2	I40474	pro-sigma-E proces
37	59.5	16.6	408	2	A99238	conserved hypothet
38	59.5	16.6	4466	1	S17653	dynein beta heavy
39	59	16.4	338	2	AG3582	iron(III)-binding
40	59	16.4	345	2	AD3024	hypothetical prote
41	59	16.4	359	2	E98260	periplasmic bindin
42	59	16.4	370	2	T34050	hypothetical prote
43	59	16.4	384	1	W2WL41	E2 protein - human
44	59	16.4	2459	2	AF2136	peptide synthetase
45	58.5	16.3	136	2	T22240	hypothetical prote

ALIGNMENTS

RESULT 1
QQCVL1
AL1 protein - tomato golden mosaic virus
C;Species: tomato golden mosaic virus
A;Note: host Nicotiana sp. (tobacco)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C;Accession: A04170
R;Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.
EMBO J. 3, 2197-2205, 1984
A;Title: Complete nucleotide sequence of the infectious cloned DNA components of toma
A;Reference number: A04163
A;Accession: A04170
A;Molecule type: DNA
A;Residues: 1-352 <HAM>
C;Comment: The genome consists of two circular, single-stranded DNA components, DNA A
.
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 95.3%; Score 342; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 2.6e-31;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 AAIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
QQCVPT
AL1 protein - potato yellow mosaic virus (isolate Venezuela)
C;Species: potato yellow mosaic virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C;Accession: JU0364
R;Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.
J. Gen. Virol. 72, 1515-1520, 1991
A;Title: The nucleotide sequence of the infectious cloned DNA components of potato ye
A;Reference number: JU0362; MUID:91311403
A;Accession: JU0364
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-361 <COU>
A;Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             67.4%;  Score 242;  DB 1;  Length 361;
  Best Local Similarity   66.7%;  Pred. No. 7e-20;
  Matches   46;  Conservative   11;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: || ||:|||||||| || |||||||||: :|| |::||:||:|||:|||:|||:  |
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 AAIFDKTPE 69
             || | ||
Db     170 DRIFMKAPE 178


RESULT   3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C;Species: cassava latent virus
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C;Accession: S07594
R;Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A;Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A;Reference number: S07590; MUID:90174930
A;Accession: S07594
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <MOR>
A;Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C;Genetics:
A;Map position: segment DNA1
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             63.8%;  Score 229;  DB 2;  Length 358;
  Best Local Similarity   60.0%;  Pred. No. 2.1e-18;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
             |: ||:||:|||||||| |::||| |:|||: || ||| :||| :|: :: |||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 AAIFDKTPEP 70
             || : | |
Db     169 DRIFQEPPAP 178


RESULT   4
JQ1887
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1887
R;Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A;Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A;Reference number: JQ1885; MUID:93139778
A;Accession: JQ1887
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-362 <DRY>
A;Cross-references: GB:S53251
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             61.8%;  Score 222;  DB 1;  Length 362;
  Best Local Similarity   52.9%;  Pred. No. 1.4e-17;
  Matches   45;  Conservative    9;  Mismatches   15;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           || |||||:|||||||| |::||| |:||| || ||| ::|| |: |: |||||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 AAI----------------FDKTPE 69
             |                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   5
JQ2300
replicase - pepper huasteco virus (component A)
N;Alternate names: ORF AL1 protein
C;Species: pepper huasteco virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C;Accession: JQ2300
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A;Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A;Reference number: JQ2299; MUID:94015007
A;Accession: JQ2300
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.2%;  Score 216;  DB 2;  Length 349;
  Best Local Similarity   58.6%;  Pred. No. 6.3e-17;
  Matches   41;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
             |: ||||||:|||||||| |::||  |:|||::| |||||||:|: |: :  ||||: |
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 AAIFDKTPEP 70
             ||   |||
Db     170 NRIFQTPPEP 179


RESULT   6
S31875
AL1 protein - pepper rizado amarillo virus
C;Species: pepper rizado amarillo virus
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Sep-1999
C;Accession: S31875
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera
submitted to the EMBL Data Library, February 1993
A;Description: Complete nucleotide sequence of pepper huasteco virus: analysis and co
A;Reference number: S31872
A;Accession: S31875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
A;Note: the source is designated as pepper huasteco virus
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.2%;  Score 216;  DB 2;  Length 349;
  Best Local Similarity   58.6%;  Pred. No. 6.3e-17;
  Matches   41;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
             |: ||||||:|||||||| |::||  |:|||::| |||||||:|: |: :  ||||: |
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 AAIFDKTPEP 70
             ||   |||
Db     170 NRIFQTPPEP 179


RESULT   7
S39211
gene C1 protein - tomato yellow leaf curl virus
```

```
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C;Accession: S39211
R;Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A;Description: High similarity among the tomato yellow leaf curl virus isolates from the
A;Reference number: S39209
A;Accession: S39211
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <NOR>
A;Cross-references: EMBL:Z25751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.9%;  Score 215;  DB 2;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 8.5e-17;
  Matches   40;  Conservative    9;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALA 61
           | || ||:|||||||| ||:||| |:|:|| || ||| :|:|  |  |:  |||:|| |
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 AIFDKTPEP 70
            :|   | |
Db     171 RVFQVPPAP 179


RESULT   8
S28360
AL1 protein - beet curly top virus
C;Species: beet curly top virus
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
C;Accession: S28360
R;Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
EMBO J. 5, 1761-1767, 1986
A;Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly to
A;Reference number: S28360
A;Accession: S28360
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-385 <STA>
A;Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.1%;  Score 212;  DB 2;  Length 385;
  Best Local Similarity   55.7%;  Pred. No. 2e-16;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: |||||:|||||||| ||:||: |:||||:| ::||||::|: |: |  | ||| :
Db     137 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 196

Qy      61 AAIFDKTPEP 70
             || : |:|
Db     197 QKIFQRPPDP 206


RESULT   9
S22593
hypothetical protein C4 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Sep-1999
C;Accession: S22593
R;Kheyr-Pour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.
Nucleic Acids Res. 19, 6763-6769, 1991
A;Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monoparti
A;Reference number: S22588; MUID:92107660
A;Accession: S22593
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-359 <KHE>
A;Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             58.8%;  Score 211;  DB 2;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 2.4e-16;
  Matches   39;  Conservative   10;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALA 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|| |
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 AIFDKTPEP 70
            :|   | |
Db     171 KVFQVPPAP 179


RESULT  10
QQCVW1
AV1 protein - abutilon mosaic virus (isolate West India)
C;Species: abutilon mosaic virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Apr-1994
C;Accession: A36214
R;Frischmuth, T.; Zimmat, G.; Jeske, H.
Virology 178, 461-468, 1990
A;Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as
A;Reference number: A36214; MUID:91020984
A;Accession: A36214
A;Molecule type: DNA
A;Residues: 1-355 <FRI>
A;Cross-references: EMBL:X15983
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             58.5%;  Score 210;  DB 1;  Length 355;
  Best Local Similarity   57.1%;  Pred. No. 3.1e-16;
  Matches   40;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |  |||||:|||||||| ||:||: |:||||   : || |::|: |: |: | ||: | |
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 AAIFDKTPEP 70
             || | |||
Db     170 ERIFAKAPEP 179


RESULT  11
JQ1870
AL1 protein - tomato mottle virus (isolate Florida)
C;Species: tomato mottle virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1870
R;Abouzid, A.M.; Polston, J.E.; Hiebert, E.
J. Gen. Virol. 73, 3225-3229, 1992
A;Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated f
A;Reference number: JQ1869; MUID:93107858
A;Accession: JQ1870
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <ABO>
A;Cross-references: GB:L14460
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             57.7%;  Score 207;  DB 1;  Length 358;
  Best Local Similarity   54.3%;  Pred. No. 6.9e-16;
  Matches   38;  Conservative   15;  Mismatches   17;  Indels    0;  Gaps    0;
```

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: ||:||:|||||||| |::||: |:|||| | : || ::||: |: :: | ||: | |
Db     107 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 166

Qy      61 AAIFDKTPEP 70
             || | |||
Db     167 ERIFAKAPEP 176
```

RESULT 12
JQ2327
AL1 protein - Indian cassava mosaic virus
N;Alternate names: replication-associated protein
C;Species: Indian cassava mosaic virus
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C;Accession: JQ2327; S35883
R;Hong, Y.G.; Robinson, D.J.; Harrison, B.D.
J. Gen. Virol. 74, 2437-2443, 1993
A;Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-tran
A;Reference number: JQ2326; MUID:94065670
A;Accession: JQ2327
A;Molecule type: DNA
A;Residues: 1-351 <HON>
A;Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             57.1%;  Score 205;  DB 2;  Length 351;
  Best Local Similarity   59.7%;  Pred. No. 1.1e-15;
  Matches   40;  Conservative    9;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALAAI 63
           || ||:|||||||| |::||| | |||: || |||:|:||  |  ||  ||:::| |  |
Db     113 WGTFQIDGRSARGGQQSANDAYAAALNSGSKSEALKILRELAPRDYLRDFHHISSNLDRI 172

Qy      64 FDKTPEP 70
           | | | |
Db     173 FTKPPPP 179
```

RESULT 13
S39235
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Sep-1997
C;Accession: S39235
R;Crespi, S.; Noris, E.; Vaira, A.; Bosco, D.; Accotto, G.
submitted to the EMBL Data Library, December 1993
A;Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.
A;Reference number: S39233
A;Accession: S39235
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <CRE>
A;Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             57.1%;  Score 205;  DB 2;  Length 359;
  Best Local Similarity   55.1%;  Pred. No. 1.2e-15;
  Matches   38;  Conservative   11;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALA 61
           | || ||:|||||||| ||:||| |:|:|| || ||| :|:: |  |:  |||::| |
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINARSKSEALDVIKQLAPRDYVLHFHNISSNLD 170

Qy      62 AIFDKTPEP 70
             :|  | |
Db     171 KVFQVPPAP 179
```

RESULT 14
S59885
replication-associated protein C1 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
C;Accession: S59885
R;Hong, Y.; Harrison, B.D.
submitted to the EMBL Data Library, February 1995
A;Description: Nucleotide sequences from tomato leaf curl viruses from different coun
d geminiviruses.
A;Reference number: S58346
A;Accession: S59885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <HON>
A;Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA88229.1; PID:g974211
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             54.3%;  Score 195;  DB 2;  Length 360;
  Best Local Similarity   57.6%;  Pred. No. 1.6e-14;
  Matches   38;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALAAI 63
           :| ||:|||||||| |::||| |||:|: ||  || |:||| |: :: ||||||: |  |
Db     113 FGVFQIDGRSARGGQQSANDAYAEAINSGSKAAALDILREKAPKDFVLQFHNLNANLDRI 172

Qy      64 FDKTPE 69
           |  : |
Db     173 FTPSAE 178
```

RESULT 15
QQCVC1
AL1 protein - tomato yellow leaf curl virus
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C;Accession: D40779
R;Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.
Virology 185, 151-161, 1991
A;Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin
A;Reference number: A40779; MUID:92024070
A;Accession: D40779
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-357 <NAV>
A;Cross-references: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             53.8%;  Score 193;  DB 1;  Length 357;
  Best Local Similarity   62.3%;  Pred. No. 2.7e-14;
  Matches   38;  Conservative   10;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALAAI 63
           :|  |:|||||||| |::||| ||||| || ||| |::|| |: |: ||||||:| |  |
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171
```

Search completed: October 10, 2002, 01:48:45
Job time: 563 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:47:18 ; Search time 133.15 Seconds
(without alignments)
58.394 Million cell updates/sec

Title: US-09-289-346A-7
Perfect score: 359
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSALAAIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	359	100.0	70	21	AAB18683	Mutant peptide der
2	342	95.3	70	21	AAB18677	Peptide fragment f
3	342	95.3	356	21	AAB18687	Amino acid sequenc
4	334	93.0	70	21	AAB18685	Mutant peptide der
5	331	92.2	70	21	AAB18688	Mutant peptide der
6	330	91.9	70	21	AAB18692	Mutant peptide der
7	328	91.4	70	21	AAB18684	Mutant peptide der
8	328	91.4	70	21	AAB18690	Mutant peptide der
9	327	91.1	70	21	AAB18678	Mutant peptide der
10	327	91.1	70	21	AAB18686	Mutant peptide der
11	326	90.8	70	21	AAB18689	Mutant peptide der
12	324	90.3	70	21	AAB18680	Mutant peptide der
13	324	90.3	70	21	AAB18691	Mutant peptide der
14	321	89.4	70	21	AAB18681	Mutant peptide der
15	318	88.6	70	21	AAB18682	Mutant peptide der
16	310	86.4	70	21	AAB18679	Mutant peptide der
17	213	59.3	353	18	AAW34338	Bean golden mosaic
18	213	59.3	353	18	AAW34332	Bean golden mosaic
19	213	59.3	353	18	AAW34333	Bean golden mosaic
20	213	59.3	353	18	AAW34334	Bean golden mosaic
21	213	59.3	353	18	AAW34335	Bean golden mosaic
22	211	58.8	353	8	AAP70407	ORF 4 gene product
23	211	58.8	359	17	AAR88870	Sardinian tomato y
24	211	58.8	359	17	AAR88871	Sardinian tomato y
25	211	58.8	359	17	AAR88872	Sardinian tomato y
26	211	58.8	361	18	AAW34336	Tomato mottle viru
27	211	58.8	361	18	AAW34324	Tomato mottle viru
28	211	58.8	361	18	AAW34325	Tomato mottle viru
29	211	58.8	361	18	AAW34326	Tomato mottle viru
30	208.5	58.1	361	8	AAP70562	Product of ORF 4 f
31	194	54.0	362	19	AAW56495	Tobacco leaf curli
32	193	53.8	357	18	AAW34329	Tomato yellow leaf
33	193	53.8	357	18	AAW34330	Tomato yellow leaf
34	193	53.8	357	18	AAW34331	Tomato yellow leaf
35	185	51.5	357	18	AAW34337	Tomato yellow leaf
36	68	18.9	447	21	AAY99659	Human GTPase assoc
37	68	18.9	447	21	AAY87089	Human secreted pro
38	68	18.9	447	21	AAY66744	Membrane-bound pro
39	68	18.9	447	21	AAY50947	Human adult aorta
40	68	18.9	447	22	AAU29150	Human PRO polypept
41	68	18.9	447	22	AAM93267	Human polypeptide,
42	68	18.9	447	22	AAE06066	Human gene 26 enco
43	68	18.9	447	22	AAB65267	Human PRO1125 (UNQ
44	68	18.9	456	21	AAY87190	Human secreted pro
45	68	18.9	456	22	AAE06167	Human gene 26 enco

ALIGNMENTS

RESULT 1
AAB18683
ID AAB18683 standard; peptide; 70 AA.
XX
AC AAB18683;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 59
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 61
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 62
FT /note= "wild type residue replaced with Ala"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.

```
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             100.0%;  Score 359;  DB 21;  Length 70;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-39;
  Matches   70;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsal 60

Qy     61 AAIFDKTPEP 70
          ||||||||||
Db     61 aaifdktpep 70


RESULT   2
AAB18677
ID   AAB18677 standard; peptide; 70 AA.
XX
AC   AAB18677;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Peptide fragment from Rep protein of TGMV (amino acids 110-179).
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Tomato golden mosaic virus.
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 18; 73pp; English.
XX
CC   The present sequence is derived from a geminivirus replication (Rep)
CC   protein, which is also known as AL1. AL1 binds double-stranded DNA,
CC   catalyses cleavage and ligation of single-stranded DNA, and interacts
CC   with other viral and host proteins. Mutants of the AL1 protein are used
CC   to produce transgenic plants. The mutation in AL1 is present in a
CC   ribosome binding region, and expression of mutant AL1 protein imparts
CC   increased resistance to geminivirus infection in the plant. Mutant AL1
CC   proteins are useful for producing plants having increased resistance or
CC   reduced sensitivity to a geminivirus such as tomato golden mosaic virus,
CC   tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl
CC   virus, African cassava mosaic virus, Indian cassava mosaic virus, potato
CC   yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,
CC   squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or
CC   beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             95.3%;  Score 342;  DB 21;  Length 70;
  Best Local Similarity   95.7%;  Pred. No. 1.2e-36;
  Matches   67;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 AAIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   3
AAB18687
ID   AAB18687 standard; peptide; 356 AA.
XX
AC   AAB18687;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Amino acid sequence of a geminivirus replication protein of TGMV.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 354
FT                   /note= "unspecified amino acid"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
```

PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT -
XX
PS Disclosure; Page 47-48; 73pp; English.
XX
CC The present sequence represents a geminivirus replication (Rep)
CC protein, which is also known as AL1. AL1 binds double-stranded DNA,
CC catalyses cleavage and ligation of single-stranded DNA, and interacts
CC with other viral and host proteins. Mutants of the AL1 protein are used
CC to produce transgenic plants. The mutation in AL1 is present in a
CC ribosome binding region, and expression of mutant AL1 protein imparts
CC increased resistance to geminivirus infection in the plant. Mutant AL1
CC proteins are useful for producing plants having increased resistance or
CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,
CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl
CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato
CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,
CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or
CC beet curly top virus.
XX
SQ Sequence 356 AA;

Query Match 95.3%; Score 342; DB 21; Length 356;
Best Local Similarity 95.7%; Pred. No. 9.3e-36;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    110 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 169

Qy     61 AAIFDKTPEP 70
          ||||||||
Db    170 drifdktpep 179
```

RESULT 4
AAB18685
ID AAB18685 standard; peptide; 70 AA.
XX
AC AAB18685;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 10
FT /note= "wild type residue replaced with Ala"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DR WPI; 2000-618851/59.
XX
PT Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT -
XX
PS Claim 53; Page 46; 73pp; English.
XX
CC The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
SQ Sequence 70 AA;

Query Match 93.0%; Score 334; DB 21; Length 70;
Best Local Similarity 94.3%; Pred. No. 1.3e-35;
Matches 66; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      1 tlvwgefqvagrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 AAIFDKTPEP 70
          ||||||||
Db     61 drifdktpep 70
```

RESULT 5
AAB18688
ID AAB18688 standard; peptide; 70 AA.
XX
AC AAB18688;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 19
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 20
FT /note= "wild type residue replaced with Ala"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DR WPI; 2000-618851/59.

```
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 48; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             92.2%;  Score 331;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 3.1e-35;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          |||||||||||||||||  |||||||||||||||||||||||||||||||||||||| |
Db      1 tlvwgefqvdgrsarggcaasndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 AAIFDKTPEP 70
          ||||||||
Db     61 drifdktpep 70


RESULT   6
AAB18692
ID   AAB18692 standard; peptide; 70 AA.
XX
AC   AAB18692;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 66
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 69
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 50; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             91.9%;  Score 330;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 4.1e-35;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 AAIFDKTPEP 70
          ||| || | |
Db     61 drifdatpap 70


RESULT   7
AAB18684
ID   AAB18684 standard; peptide; 70 AA.
XX
AC   AAB18684;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 7
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 8
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
```

```
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             91.4%;  Score 328;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 7.5e-35;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          ||||||  |||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      1 tlvwgeaavdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 AAIFDKTPEP 70
              ||||||||
Db     61 drifdktpep 70


RESULT   8
AAB18690
ID   AAB18690 standard; peptide; 70 AA.
XX
AC   AAB18690;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 27
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 30
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             91.4%;  Score 328;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 7.5e-35;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          |||||||||||||||||||||||||| || |||||||||||||||||||||||||| |
Db      1 tlvwgefqvdgrsarggcqtsndaaaaalaasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 AAIFDKTPEP 70
              ||||||||
Db     61 drifdktpep 70


RESULT   9
AAB18678
ID   AAB18678 standard; peptide; 70 AA.
XX
AC   AAB18678;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 12
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 13
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 15
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
```

```
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 42-43; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             91.1%;  Score 327;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 1e-34;
  Matches   64;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |||||||||| :| ||||||||||||||||||||||||||||||||||||||||||| |
Db       1 tlvwgefqvdgaaaaggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 AAIFDKTPEP 70
           ||||||||
Db      61 drifdktpep 70


RESULT  10
AAB18686
ID   AAB18686 standard; peptide; 70 AA.
XX
AC   AAB18686;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 24
FT                   /note= "wild type residue replaced with Leu"
FT   Misc-difference 25
FT                   /note= "wild type residue replaced with Leu"
FT   Misc-difference 26
FT                   /note= "wild type residue replaced with Leu"
XX
```

```
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 46; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             91.1%;  Score 327;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 1e-34;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |||||||||||||||||||||   |||||||||||||||||||||||||||||||| |
Db       1 tlvwgefqvdgrsarggcqtsndlllealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 AAIFDKTPEP 70
           ||||||||
Db      61 drifdktpep 70


RESULT  11
AAB18689
ID   AAB18689 standard; peptide; 70 AA.
XX
AC   AAB18689;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 22
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 23
```

```
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 48-49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             90.8%;  Score 326;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 1.4e-34;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |||||||||||||||||||  |||||||||||||||||||||||||||||||| |
Db       1 tlvwgefqvdgrsarggcqtsaaaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 AAIFDKTPEP 70
           ||||||||
Db      61 drifdktpep 70


RESULT  12
AAB18680
ID   AAB18680 standard; peptide; 70 AA.
XX
AC   AAB18680;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 42
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 43
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 44
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 43-44; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             90.3%;  Score 324;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 2.5e-34;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |||||||||||||||||||||||||||||||||||||||||||   |||||||||||||| |
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiiaaaipekylfqfhnlnsnl 60

Qy      61 AAIFDKTPEP 70
           ||||||||
Db      61 drifdktpep 70


RESULT  13
AAB18691
ID   AAB18691 standard; peptide; 70 AA.
XX
AC   AAB18691;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
```

```
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 34
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 35
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 36
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match           90.3%;  Score 324;  DB 21;  Length 70;
  Best Local Similarity  91.4%;  Pred. No. 2.5e-34;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          ||||||||||||||||||||||||||||||||   |||||||||||||||||||| |
Db      1 tlvwgefqvdgrsarggcqtsndaaaealnassaaaalqiirekipekylfqfhnlnsnl 60

Qy     61 AAIFDKTPEP 70
          ||||||||
Db     61 drifdktpep 70


RESULT  14
AAB18681
ID   AAB18681 standard; peptide; 70 AA.
XX
AC   AAB18681;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 47
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 48
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 49
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 44; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match           89.4%;  Score 321;  DB 21;  Length 70;
  Best Local Similarity  91.4%;  Pred. No. 6e-34;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||   |||||||| |
Db      1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipaaalfqfhnlnsnl 60

Qy     61 AAIFDKTPEP 70
          ||||||||
Db     61 drifdktpep 70


RESULT  15
AAB18682
ID   AAB18682 standard; peptide; 70 AA.
XX
AC   AAB18682;
XX
```

```
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 52
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 54
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 55
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 44-45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             88.6%;  Score 318;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 1.5e-33;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||| |  ||| |
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfafaalnsnl 60

Qy      61 AAIFDKTPEP 70
             ||||||||
Db      61 drifdktpep 70


Search completed: October 10, 2002, 01:47:19

Job time: 528 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:02:00 ; Search time 118.38 Seconds
(without alignments)
102.295 Million cell updates/sec

Title: US-09-289-346A-8
Perfect score: 361
Sequence: 1 TLVWGEAAVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	301	83.4	351	12	Q91R10	Q91r10 tomato seve
2	299	82.8	352	12	Q9E000	Q9e000 tomato rugo
3	282	78.1	226	12	O09727	O09727 leonurus mo
4	282	78.1	226	12	Q9WHF6	Q9whf6 tomato mild
5	280	77.6	361	12	Q67574	Q67574 bean golden
6	278	77.0	225	12	Q9QDB1	Q9qdb1 cowpea gold
7	278	77.0	314	12	Q9ELT8	Q9elt8 sweet potat
8	278	77.0	364	12	Q9QS55	Q9qs55 sweet potat
9	275	76.2	185	12	Q98693	Q98693 sida golden
10	267	74.0	149	12	P88975	P88975 macroptiliu
11	267	74.0	233	12	Q9YLA4	Q9yla4 macroptiliu
12	261	72.3	234	12	O39180	O39180 geminivirid
13	258	71.5	190	12	Q9Z089	Q9z089 tobacco lea
14	258	71.5	190	12	Q9Z084	Q9z084 tobacco lea
15	257	71.2	190	12	Q9W827	Q9w827 tobacco lea
16	257	71.2	208	12	Q9Z0C4	Q9z0c4 tobacco lea
17	257	71.2	359	12	Q91M88	Q91m88 tobacco lea
18	255	70.6	208	12	Q9Z0C0	Q9z0c0 tobacco lea
19	255	70.6	208	12	Q9Z0B8	Q9z0b8 tobacco lea
20	253	70.1	208	12	Q9Z0B6	Q9z0b6 tobacco lea
21	252	69.8	203	12	Q9Z083	Q9z083 tobacco lea
22	252	69.8	363	12	O73577	O73577 cotton leaf
23	247	68.4	359	12	Q91B86	Q91b86 ageratum ye
24	247	68.4	360	12	Q9DXI0	Q9dxi0 ageratum ye
25	245	67.9	208	12	Q9Z0A0	Q9z0a0 tobacco lea
26	245	67.9	363	12	O72705	O72705 cotton leaf
27	245	67.9	363	12	O72719	O72719 cotton leaf
28	244	67.6	190	12	Q9Z086	Q9z086 tobacco lea
29	244	67.6	208	12	Q9Z0C6	Q9z0c6 tobacco lea
30	243	67.3	349	12	Q88888	Q88888 tomato pseu
31	243	67.3	362	12	O56816	O56816 chayote mos
32	242	67.0	359	12	Q9YZV4	Q9yzv4 tomato yell
33	242	67.0	359	12	Q9YZV2	Q9yzv2 tomato yell
34	242	67.0	359	12	Q9YUX7	Q9yux7 tomato yell
35	242	67.0	359	12	Q9YL27	Q9yl27 tomato yell
36	241	66.8	348	12	Q911W5	Q911w5 macroptiliu
37	241	66.8	353	12	O72692	O72692 beet curly
38	241	66.8	354	12	Q65438	Q65438 beet curly
39	241	66.8	359	12	Q88942	Q88942 tomato yell
40	240	66.5	190	12	Q9Z0A7	Q9z0a7 tobacco lea
41	240	66.5	362	12	Q9IN48	Q9in48 okra enatio
42	239	66.2	307	12	Q9IET7	Q9iet7 cotton leaf
43	239	66.2	361	12	O72723	O72723 cotton leaf
44	238	65.9	231	12	Q96620	Q96620 african tom
45	238	65.9	354	12	Q9IN42	Q9in42 south afric

ALIGNMENTS

RESULT 1
Q91R10
ID Q91R10 PRELIMINARY; PRT; 351 AA.
AC Q91R10;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE REP PROTEIN.
GN AC1.
OS Tomato severe rugose virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=158463;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MINAS GERAIS;
RA Rezende W.L., Goulart L.R., Parreira K.S., Figueiredo J.E.F.;
RT "The full-length DNA-A nucleotide sequence of a novel tomato-infecting
RT begomovirus, Tomato severe rugose virus, in Brazil.";
RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY029750; AAK50357.1; -.
SQ SEQUENCE 351 AA; 40122 MW; 87F937A4F873B6CF CRC64;

Query Match 83.4%; Score 301; DB 12; Length 351;
Best Local Similarity 80.0%; Pred. No. 5.6e-26;
Matches 56; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

```
Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||  :||||||||||:|||||||||| ||: ||||||||:|||:||||||||||
Db     111 TIEWGEFQIDGRSARGGCQTANDAAAEALNAPSKDVALQIIREKLPEKFLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| : |||
Db     171 DRIFARAPEP 180
```

RESULT 2
Q9E000
ID Q9E000 PRELIMINARY; PRT; 352 AA.

```
AC   Q9E000;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN.
GN   AC1.
OS   Tomato rugose mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=134599;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Fernandes J.J., Fontes E.P.B., Brommonschenkel S.H., Carvalho M.G.,
RA   Zambolim E.M., Zerbini F.M.;
RT   "Molecular Cloning and Characterization of Tomato rugose mosaic virus
RT   (TRMV), a Begomovirus Isolated from Tomatoes at Triangulo Mineiro,
RT   Minas Gerais, Brazil.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF291705; AAG15546.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   352 AA;  40012 MW;  47CD55838E24D613 CRC64;


  Query Match             82.8%;  Score 299;  DB 12;  Length 352;
  Best Local Similarity   80.0%;  Pred. No. 9.5e-26;
  Matches   56;  Conservative    7;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||  :|||||||||||:|||||||||| ||: |||||||:|||:||||||||||
Db     111 TIEWGEFQIDGRSARGGCQTANDAAAEALNAPSKDIALQIIREKMPEKFLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| : |||
Db     171 DRIFARAPEP 180


RESULT   3
O09727
ID   O09727      PRELIMINARY;      PRT;   226 AA.
AC   O09727;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   REP.
OS   Leonurus mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=58177;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LEMV- BRAZIL 1;
RA   Faria J.C., Maxwell D.P.;
RT   "Variability in geminivirus associated with Phaseolus vulgaris in
RT   Brazil.";
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U92532; AAB51157.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     226    226
SQ   SEQUENCE   226 AA;  25617 MW;  73CDB6E766083FC5 CRC64;


  Query Match             78.1%;  Score 282;  DB 12;  Length 226;
  Best Local Similarity   78.6%;  Pred. No. 4.8e-24;
  Matches   55;  Conservative    4;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||  ||||:||| || ||||||||||  |  |||||:||:||||||||||||||||
Db     111 TVEWGEFQVDGRSSRGGQQTVNDAAAEALNAPDKRTALQIIKEKLPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180


RESULT   4
Q9WHF6
ID   Q9WHF6      PRELIMINARY;      PRT;   226 AA.
AC   Q9WHF6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   tomato mild mottle geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX   NCBI_TaxID=92943;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HN96-H5KW;
RA   Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA   Maxwell D.P.;
RT   "Molecular characterization and DNA-based detection methods for
RT   vegetable-infecting geminiviruses in Central America.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF131071; AAD33471.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     226    226
SQ   SEQUENCE   226 AA;  25941 MW;  2EA4116712871A23 CRC64;


  Query Match             78.1%;  Score 282;  DB 12;  Length 226;
  Best Local Similarity   74.3%;  Pred. No. 4.8e-24;
  Matches   52;  Conservative   12;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||  :|||||||| ||:||||||||||||||||::||:||:|:::|||:|||:|||
Db     111 TIEWGEFQIDGRSARGGQQTANDAAAEALNASSKEEAMRIIKEKLPKEFLFQYHNLSSNL 170

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180


RESULT   5
Q67574
ID   Q67574      PRELIMINARY;      PRT;   361 AA.
AC   Q67574;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE  REPLICATIVE PROTEIN.
GN   AL1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,
RA   Maxwell D.P., Russell D.R.;
RT   "Cloning of the complete DNA genomes of four bean-infecting
RT   geminiviruses and determining their infectivity by electric discharge
RT   particle acceleration.";
RL   Phytopathology 81:980-985(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,
RA   Morales F.J., Maxwell D.P.;
RT   "Differentiation of bean-infecting geminiviruses by nucleic acid
```

```
RT   hybridization probes and aspects of bean golden mosaic in Brazil.";
RL   Plant Dis. 75:336-342(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RT   "Genetic diversity in geminiviruses causing bean golden mosaic
RT   disease: The nucleotide sequence of the infectious cloned DNA
RT   components of a Brazilian isolate of bean golden mosaic geminivirus.";
RL   Submitted (MAY-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; M88686; AAA46312.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   361 AA;  41041 MW;  0094C7ACAF06B788 CRC64;


  Query Match             77.6%;  Score 280;  DB 12;  Length 361;
  Best Local Similarity   77.6%;  Pred. No. 1.4e-23;
  Matches   52;  Conservative    9;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       4 WGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           ||   ||||||||| ||:||||:||||||||||||:|||:||:|||:|||:|||:||||||
Db     113 WGHFQVDGRSARGGQQTANDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
           | | |:|
Db     173 FTKAPDP 179


RESULT   6
Q9QDB1
ID   Q9QDB1      PRELIMINARY;      PRT;   225 AA.
AC   Q9QDB1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   cowpea golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=69263;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CGMV-BR;
RA   Faria J.C.;
RT   "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT   Brazil.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF188708; AAF06318.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     225     225
SQ   SEQUENCE   225 AA;  25766 MW;  1089CB6BD8D15B5D CRC64;


  Query Match             77.0%;  Score 278;  DB 12;  Length 225;
  Best Local Similarity   77.6%;  Pred. No. 1.3e-23;
  Matches   52;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       4 WGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           ||   :|||||||| || ||||:||||||||||||:|||:|||:|||:|||:||||||
Db     113 WGHFQIDGRSARGGQQTINDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
           | | |||
Db     173 FKKPPEP 179


RESULT   7
Q9ELT8
ID   Q9ELT8      PRELIMINARY;      PRT;   314 AA.
AC   Q9ELT8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION ASSOCIATION PROTEIN.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF288227; AAG01006.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   314 AA;  35153 MW;  6B6220613046943F CRC64;


  Query Match             77.0%;  Score 278;  DB 12;  Length 314;
  Best Local Similarity   80.9%;  Pred. No. 2e-23;
  Matches   55;  Conservative    3;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||  ||||||||| ||:|||||||||| ||| ||||||||:|||||||||| |||
Db     110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy      61 DRIFDKTP 68
           ||||   |
Db     170 DRIFSPPP 177


RESULT   8
Q9QS55
ID   Q9QS55      PRELIMINARY;      PRT;   364 AA.
AC   Q9QS55;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION INITIATION PROTEIN AC1.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A.;
RT   "Cloning of a DNA-A-like genomic component of sweet potato leaf curl
RT   virus : nucleotide sequence and phylogenetic relationships.";
RL   Online Publication.
DR   EMBL; AF104036; AAD47173.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   364 AA;  40680 MW;  5F79752431A09D6E CRC64;
```

```
  Query Match             77.0%;  Score 278;  DB 12;  Length 364;
  Best Local Similarity   80.9%;  Pred. No. 2.3e-23;
  Matches   55;  Conservative    3;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||  |||||||||| ||:||||||||||| ||| ||||||||:||||||||||| |||
Db    110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy     61 DRIFDKTP 68
          ||||   |
Db    170 DRIFSPPP 177


RESULT   9
Q98693
ID   Q98693      PRELIMINARY;      PRT;   185 AA.
AC   Q98693;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   AC1.
OS   sida golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51034;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA;
RA   Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RT   "Genetic Diversity among geminiviruses associated with the weed
RT   species Sida spp, Macroptilium lathyroides, and Wissadula amplissima
RT   from Jamaica.";
RL   Plant Dis. 81:1251-1258(1997).
DR   EMBL; U67926; AAB97865.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     185    185
SQ   SEQUENCE   185 AA;  20975 MW;  3913850A025A5EE1 CRC64;


  Query Match             76.2%;  Score 275;  DB 12;  Length 185;
  Best Local Similarity   71.4%;  Pred. No. 2.3e-23;
  Matches   50;  Conservative   12;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||   :||||||||| ||:||||||||: :||:||:|||||:||:||||:|||:||:
Db     89 TIEWGVFQIDGRSARGGQQTANDAAAEALNSGTKEDALKIIREKLPERYLFQYHNLSSNI 148

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    149 DRIFSKPPEP 158


RESULT  10
P88975
ID   P88975      PRELIMINARY;      PRT;   149 AA.
AC   P88975;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   AC1.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICAN;
RA   Roye M.E.;
RL   Thesis (1996), Biochemistry, University of the West Indies, Jamaica.
DR   EMBL; U75278; AAB36919.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     149    149
SQ   SEQUENCE   149 AA;  16785 MW;  E4CF5EED4C9CD508 CRC64;


  Query Match             74.0%;  Score 267;  DB 12;  Length 149;
  Best Local Similarity   70.0%;  Pred. No. 1.5e-22;
  Matches   49;  Conservative   12;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||   :|||||||| ||||||||||||: :|| |::|::||:|||:|||:|||:|||
Db     52 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 111

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    112 DRIFMKDPEP 121


RESULT  11
Q9YLA4
ID   Q9YLA4      PRELIMINARY;      PRT;   233 AA.
AC   Q9YLA4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E.;
RT   "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
RT   from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E., McLaughlin W.A., Maxwell D.P.;
RT   "Molecular characterization of two distinct geminiviruses infecting M.
RT   lathyroides from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF098940; AAD17850.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     233    233
SQ   SEQUENCE   233 AA;  26356 MW;  AA490AF4D2166A02 CRC64;


  Query Match             74.0%;  Score 267;  DB 12;  Length 233;
  Best Local Similarity   70.0%;  Pred. No. 2.5e-22;
  Matches   49;  Conservative   12;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||   :|||||||| ||||||||||||: :|| |::|::||:|||:|||:|||:|||
Db    110 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 169

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    170 DRIFMKDPEP 179


RESULT  12
O39180
```

```
ID   O39180      PRELIMINARY;      PRT;   234 AA.
AC   O39180;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   AC1 (FRAGMENT).
OS   Geminiviridae.
OC   Viruses; ssDNA viruses.
OX   NCBI_TaxID=10811;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOMATO STRAIN;
RA   Guzman P., Arredondo C.R., Emmatty D., Portillo R.J., Gilbertson R.L.;
RT   "Partial Characterization of Two Whitefly-Transmitted Geminiviruses
RT   Infecting Tomatoes in Venezuela.";
RL   Plant Dis. 81:312-312(1997).
DR   EMBL; AF026553; AAB82605.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     234    234
SQ   SEQUENCE   234 AA;  26486 MW;  9ED8F0697105CD19 CRC64;


  Query Match             72.3%;  Score 261;  DB 12;  Length 234;
  Best Local Similarity   68.6%;  Pred. No. 1.2e-21;
  Matches   48;  Conservative   13;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||   :|||||||| ||:|||||||||: :||||::|||:||:|||:|||:|||:|:|
Db     110 TIEWGLFQIDGRSARGGQQTANDAAAEALNSGTKEEAMKIIKEKLPEKFLFQYHNLSSDL 169

Qy      61 DRIFDKTPEP 70
           |||| | | |
Db     170 DRIFMKAPNP 179


RESULT  13
Q9Z089
ID   Q9Z089      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z089;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA3-1, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001315; BAA34033.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21432 MW;  AAC093D1D1610FAD CRC64;


  Query Match             71.5%;  Score 258;  DB 12;  Length 190;
  Best Local Similarity   60.0%;  Pred. No. 2e-21;
  Matches   51;  Conservative    8;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || ||   :|||||||||| :||| |||||||| ||| |||||:|: ::||:|||||||
Db     85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db    145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  14
Q9Z084
ID   Q9Z084      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z084;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA5-2, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA5;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001318; BAA34039.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  AAC1C2943E3F01AD CRC64;


  Query Match             71.5%;  Score 258;  DB 12;  Length 190;
  Best Local Similarity   60.0%;  Pred. No. 2e-21;
  Matches   51;  Conservative    8;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || ||   :|||||||||| :||| |||||||| ||| |||||:|: ::||:|||||||
Db     85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db    145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  15
Q9W827
ID   Q9W827      PRELIMINARY;      PRT;   190 AA.
AC   Q9W827;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE GORA1-1, PARTIAL AND
DE   COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GORA1;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001303; BAA34010.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
```

```
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  93C3742A8EBDB7EB CRC64;


  Query Match             71.2%;  Score 257;  DB 12;  Length 190;
  Best Local Similarity   60.0%;  Pred. No. 2.6e-21;
  Matches   51;  Conservative    8;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || ||   |||||||||||| :||| |||||||:|| ||| ||||||:|: ::||:|||||||
Db      85 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFSNSSFDQVPE 169


Search completed: October 10, 2002, 02:02:00
Job time: 1093 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:09:52 ; Search time 34.62 Seconds
(without alignments)
78.289 Million cell updates/sec

Title: US-09-289-346A-8
Perfect score: 361
Sequence: 1 TLVWGEAAVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	350	97.0	352	1	VAL1_TGMV	P03567 tomato gold
2	250	69.3	361	1	VAL1_PYMVV	P27258 potato yell
3	237	65.7	358	1	VAL1_CLVK	P14982 cassava lat
4	237	65.7	358	1	VAL1_CLVN	P14972 cassava lat
5	230	63.7	362	1	VAL1_TYLCA	P36279 tomato yell
6	223	61.8	359	1	VAL1_TYLCU	P38609 tomato yell
7	219	60.7	349	1	VAL1_PHUV	Q06923 pepper huas
8	216	59.8	359	1	VAL1_TYLCM	P27260 tomato yell
9	214	59.3	353	1	VAL1_BGMV	P05175 bean golden
10	213	59.0	355	1	VAL1_ABMVW	P21947 abutilon mo
11	212	58.7	357	1	VAL1_TYLCV	P27259 tomato yell
12	210	58.2	358	1	VAL1_BCTV	P14991 beet curly
13	210	58.2	361	1	VAL1_TMOV	Q06657 tomato mott
14	117	32.4	347	1	VAL1_SLCV	P29048 squash leaf
15	68.5	19.0	1713	1	LMA3_HUMAN	Q16787 homo sapien
16	64.5	17.9	299	1	Y175_HELPJ	Q9zmq7 helicobacte
17	63.5	17.6	447	1	TBL2_HUMAN	Q9y4p3 homo sapien
18	63.5	17.6	1610	1	CCAD_MESAU	Q99244 mesocricetu
19	63.5	17.6	2161	1	CCAD_HUMAN	Q01668 homo sapien
20	63.5	17.6	2203	1	CCAD_RAT	P27732 rattus norv
21	62	17.2	295	1	VAL1_TYDVA	P31617 tobacco yel
22	61	16.9	1483	1	CYP1_YEAST	P12351 saccharomyc
23	60	16.6	432	1	PROA_DEIRA	Q9rtd9 deinococcus
24	60	16.6	447	1	CDSA_DROME	P56079 d phosphati
25	59.5	16.5	136	1	Y452_CAEEL	O62250 caenorhabdi
26	59.5	16.5	299	1	Y175_HELPY	P56112 helicobacte
27	59	16.3	640	1	DNAK_RHOMR	Q9xcb1 rhodothermu
28	58.5	16.2	297	1	RRPP_RABVP	P06747 rabies viru
29	58.5	16.2	334	1	G3P_BACST	P00362 bacillus st
30	58.5	16.2	335	1	G3P_BACCO	P15115 bacillus co
31	58.5	16.2	428	1	GBA1_CANAL	P28868 candida alb
32	58.5	16.2	589	1	YE67_HAEIN	P45221 haemophilus
33	58.5	16.2	1044	1	BUB1_SCHPO	O94751 schizosacch
34	58.5	16.2	1070	1	P11B_HUMAN	P42338 homo sapien
35	58	16.1	319	1	LDH2_STAAM	Q99r35 staphylococ
36	58	16.1	617	1	YACH_ECOLI	P36682 escherichia
37	58	16.1	874	1	SLAP_BACLI	P49052 bacillus li
38	57.5	15.9	439	1	XYLA_LACLA	Q9cfg7 lactococcus
39	57.5	15.9	925	1	Y562_HUMAN	O60308 homo sapien
40	57	15.8	200	1	VIP_CHICK	P48143 gallus gall
41	57	15.8	200	1	VIP_MELGA	P45644 meleagris g
42	57	15.8	247	1	YCP4_YEAST	P25349 saccharomyc
43	57	15.8	355	1	CRTB_RHOSH	P54905 rhodobacter
44	57	15.8	387	1	Y4PF_RHISN	P55615 rhizobium s
45	57	15.8	4639	1	DYHC_DROME	P37276 drosophila

ALIGNMENTS

RESULT 1
VAL1_TGMV
ID VAL1_TGMV STANDARD; PRT; 352 AA.
AC P03567;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein.
GN AC1.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10831;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;
RT "Complete nucleotide sequence of the infectious cloned DNA components
RT of tomato golden mosaic virus: potential coding regions and regulatory
RT sequences.";
RL EMBO J. 3:2197-2205(1984).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; K02029; -; NOT_ANNOTATED_CDS.
DR PIR; A04170; QQCVL1.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR ProDom; PD000736; Gemini_AL1; 1.
KW ATP-binding.
FT NP_BIND 223 230 ATP (POTENTIAL).
SQ SEQUENCE 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;

Query Match 97.0%; Score 350; DB 1; Length 352;
Best Local Similarity 97.1%; Pred. No. 4.6e-32;
Matches 68; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
VAL1_PYMVV
ID VAL1_PYMVV STANDARD; PRT; 361 AA.

```
AC   P27258;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 protein.
OS   Potato yellow mosaic virus (isolate Venezuela).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10828;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91311403; PubMed=1856690;
RA   Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT   "The nucleotide sequence of the infectious cloned DNA components of
RT   potato yellow mosaic virus.";
RL   J. Gen. Virol. 72:1515-1520(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; D00940; BAA00782.1; -.
DR   PIR; JU0364; QQCVPT.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   361 AA;  40850 MW;  5627A33BF1264383 CRC64;


  Query Match             69.3%;  Score 250;  DB 1;  Length 361;
  Best Local Similarity   68.1%;  Pred. No. 8.4e-21;
  Matches   47;  Conservative   11;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||   :||||||||| || |||||||||: :|| |::||:||:|||:|||:|||: ||
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 DRIFDKTPE 69
           |||| | ||
Db     170 DRIFMKAPE 178


RESULT   3
VAL1_CLVK
ID   VAL1_CLVK      STANDARD;      PRT;   358 AA.
AC   P14982;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (40.4 kDa protein).
GN   AC1.
OS   Cassava latent virus (strain West Kenyan 844).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10818;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Gay M.R.;
RT   "Nucleotide sequence of cassava latent virus DNA.";
RL   Nature 301:260-262(1983).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; J02057; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40346 MW;  ED173E753EE92D69 CRC64;


  Query Match             65.7%;  Score 237;  DB 1;  Length 358;
  Best Local Similarity   61.4%;  Pred. No. 2.4e-19;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:  :|||||||| |::||| |:|||: || ||| :||| :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   4
VAL1_CLVN
ID   VAL1_CLVN      STANDARD;      PRT;   358 AA.
AC   P14972;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (40.4 kDa protein).
GN   AC1.
OS   Cassava latent virus (strain Nigerian).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10819;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90174930; PubMed=2308831;
RA   Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT   "Nucleotide sequence of the infectious cloned DNA components of
RT   African cassava mosaic virus (Nigerian strain).";
RL   Nucleic Acids Res. 18:197-198(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X17095; CAA34953.1; -.
DR   PIR; S07594; S07594.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40435 MW;  1DB16BB0CB2D5E2C CRC64;


  Query Match             65.7%;  Score 237;  DB 1;  Length 358;
  Best Local Similarity   61.4%;  Pred. No. 2.4e-19;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:  :|||||||| |::||| |:|||: || ||| :||| :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168
```

```
Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   5
VAL1_TYLCA
ID   VAL1_TYLCA     STANDARD;      PRT;   362 AA.
AC   P36279;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=36447;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93139778; PubMed=8423446;
RA   Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;
RT   "Nucleotide sequence and genome organization of tomato leaf curl
RT   geminivirus.";
RL   J. Gen. Virol. 74:147-151(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
DR   PIR; JQ1887; JQ1887.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   362 AA;  41197 MW;  343E7184B4704098 CRC64;


  Query Match             63.7%;  Score 230;  DB 1;  Length 362;
  Best Local Similarity   54.1%;  Pred. No. 1.5e-18;
  Matches   46;  Conservative    9;  Mismatches   14;  Indels   16;  Gaps    1;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || |||  :|||||||| |::||| |:|||  || ||| ::||  |: |: |||||||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   6
VAL1_TYLCU
ID   VAL1_TYLCU     STANDARD;      PRT;   359 AA.
AC   P38609;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=37139;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94256836; PubMed=8198442;
RA   Noris E., Hidalgo E., Accotto G., Moriones E.;
RT   "High similarity among the tomato yellow leaf curl virus isolates
RT   from the west Mediterranean basin: the nucleotide sequence of an
RT   infectious clone from Spain.";
RL   Arch. Virol. 135:165-170(1994).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z25751; CAA81026.1; -.
DR   PIR; S39211; S39211.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  41065 MW;  2D170A51EF80A3EC CRC64;


  Query Match             61.8%;  Score 223;  DB 1;  Length 359;
  Best Local Similarity   59.4%;  Pred. No. 9.1e-18;
  Matches   41;  Conservative    9;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | ||   :|||||||| ||:||| |:|:|| || ||| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   7
VAL1_PHUV
ID   VAL1_PHUV      STANDARD;      PRT;   349 AA.
AC   Q06923;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AL1.
OS   Pepper huasteco virus (PHV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=28349;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94015007; PubMed=8409944;
RA   Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA   Rivera-Bustamante R.F.;
RT   "Complete nucleotide sequence of pepper huasteco virus: analysis and
RT   comparison with bipartite geminiviruses.";
RL   J. Gen. Virol. 74:2225-2231(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70418; CAA49856.1; -.
DR   PIR; S31875; S31875.
DR   PIR; JQ2300; JQ2300.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (BY SIMILARITY).
SQ   SEQUENCE   349 AA;  39722 MW;  D5F4E76CD56370F4 CRC64;


  Query Match             60.7%;  Score 219;  DB 1;  Length 349;
```

```
  Best Local Similarity   58.6%;  Pred. No. 2.5e-17;
  Matches   41;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||  :|||||||| |::||  |:|||::| |||||||:|: |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   8
VAL1_TYLCM
ID   VAL1_TYLCM     STANDARD;      PRT;   359 AA.
AC   P27260;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10833;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92107660; PubMed=1840676;
RA   Kheyr-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,
RA   Gronenborn B.;
RT   "Tomato yellow leaf curl virus from Sardinia is a
RT   whitefly-transmitted monopartite geminivirus.";
RL   Nucleic Acids Res. 19:6763-6769(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X61153; CAA43466.1; -.
DR   PIR; S22593; S22593.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  40733 MW;  9717B4A07C93EFA7 CRC64;


  Query Match             59.8%;  Score 216;  DB 1;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 5.6e-17;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | ||   :|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   9
VAL1_BGMV
ID   VAL1_BGMV      STANDARD;      PRT;   353 AA.
AC   P05175;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   AL1 protein (40.2 kDa protein).
GN   AC1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Howarth A.J., Caton J., Bossert M., Goodman R.M.;
RT   "Nucleotide sequence of bean golden mosaic virus and a model for gene
RT   regulation in geminiviruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M10070; AAA46318.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   353 AA;  40190 MW;  80FA779DF6029A34 CRC64;


  Query Match             59.3%;  Score 214;  DB 1;  Length 353;
  Best Local Similarity   58.6%;  Pred. No. 9.2e-17;
  Matches   41;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:  |||||||||| |::||: |:|||| | | || |::|: |: |: | ||: |||
Db     110 TIEWGQFQVDGRSARGGQQSANDSYAKALNADSIESALTILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFFKVPEP 179


RESULT  10
VAL1_ABMVW
ID   VAL1_ABMVW     STANDARD;      PRT;   355 AA.
AC   P21947;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AC1.
OS   Abutilon mosaic virus (isolate West India).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10816;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91020984; PubMed=2219703;
RA   Frischmuth T., Zimmat G., Jeske H.;
RT   "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT   as well as eukaryotic features.";
RL   Virology 178:461-468(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR   PIR; A36214; QQCVW1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   355 AA;  40257 MW;  16A2CA8A63251E95 CRC64;


  Query Match             59.0%;  Score 213;  DB 1;  Length 355;
  Best Local Similarity   57.1%;  Pred. No. 1.2e-16;
  Matches   40;  Conservative   12;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |  |||  :|||||||| ||:||: |:||||   : || |::|: |: |: | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
VAL1_TYLCV
ID   VAL1_TYLCV     STANDARD;      PRT;   357 AA.
AC   P27259;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10832;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92024070; PubMed=1926771;
RA   Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;
RT   "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RT   with a single genomic component.";
RL   Virology 185:151-161(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15656; CAA33688.1; -.
DR   PIR; D40779; QQCVC1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     219    226       ATP (POTENTIAL).
SQ   SEQUENCE   357 AA;  40678 MW;  939AB68E1AB3B2A7 CRC64;


  Query Match             58.7%;  Score 212;  DB 1;  Length 357;
  Best Local Similarity   65.6%;  Pred. No. 1.6e-16;
  Matches   40;  Conservative   11;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       4 WGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           :| : :|||||||| |::||| ||||||: || ||| |::|| |: |: ||||||:||||||
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171


RESULT  12
VAL1_BCTV
ID   VAL1_BCTV      STANDARD;      PRT;   358 AA.
AC   P14991;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 protein (40.8 kDa protein).
OS   Beet curly top virus (BCTV).
OC   Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
OX   NCBI_TaxID=10840;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT   "The nucleotide sequence of an infectious clone of the geminivirus
RT   beet curly top virus.";
RL   EMBO J. 5:1761-1767(1986).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40889 MW;  39A45FE3C0B9C333 CRC64;


  Query Match             58.2%;  Score 210;  DB 1;  Length 358;
  Best Local Similarity   54.3%;  Pred. No. 2.6e-16;
  Matches   38;  Conservative   16;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||  :|||||||| ||:||: |:||||:| ::||||::|: |: |  | ||| :|
Db     110 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 169

Qy      61 DRIFDKTPEP 70
           :|| : |:|
Db     170 QKIFQRPPDP 179


RESULT  13
VAL1_TMOV
ID   VAL1_TMOV      STANDARD;      PRT;   361 AA.
AC   Q06657;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AL1.
OS   Tomato mottle virus (isolate Florida) (TMoV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=36449;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93107858; PubMed=1469361;
RA   Abouzid A.M., Polston J.E., Hiebert E.;
RT   "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT   isolated from tomatoes in Florida.";
RL   J. Gen. Virol. 73:3225-3229(1992).
```

```
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L14460; AAC32414.1; -.
DR   PIR; JQ1870; JQ1870.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (BY SIMILARITY).
SQ   SEQUENCE   361 AA;  40516 MW;  8138B65CEEAC6950 CRC64;


  Query Match             58.2%;  Score 210;  DB 1;  Length 361;
  Best Local Similarity   54.3%;  Pred. No. 2.7e-16;
  Matches   38;  Conservative   16;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:  :|||||||| |::||: |:|||| | : || ::||: |: :: | ||: |||
Db     110 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
             :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  14
VAL1_SLCV
ID   VAL1_SLCV      STANDARD;      PRT;   347 AA.
AC   P29048;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   AL1 protein.
OS   Squash leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10829;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91082449; PubMed=1984668;
RA   Lazarowitz S.G., Lazdins I.B.;
RT   "Infectivity and complete nucleotide sequence of the cloned genomic
RT   components of a bipartite squash leaf curl geminivirus with a broad
RT   host range phenotype.";
RL   Virology 180:58-69(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M38183; AAC32410.1; ALT_INIT.
DR   PIR; C36785; QQCVS1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     218    225       ATP (POTENTIAL).
SQ   SEQUENCE   347 AA;  39110 MW;  AFDAEBDDE122110E CRC64;


  Query Match             32.4%;  Score 117;  DB 1;  Length 347;
  Best Local Similarity   37.9%;  Pred. No. 7.3e-06;
  Matches   25;  Conservative   12;  Mismatches   25;  Indels    4;  Gaps    1;

Qy       5 GEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRIF 64
           |:  | |    |      :|    |:|| |  ||| ||:   |: ::  :||| :|::|:|
Db     116 GQYKVSG----GSKSNKDDVYHNAVNAGSAGEALDIIKAGDPKTFIVNYHNLLANVERLF 171

Qy      65 DKTPEP 70
             | |||
Db     172 QKPPEP 177


RESULT  15
LMA3_HUMAN
ID   LMA3_HUMAN     STANDARD;      PRT;  1713 AA.
AC   Q16787; Q13679; Q13680;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Laminin alpha-3 chain precursor (Epiligrin 170 kDa subunit) (E170)
DE   (Nicein alpha subunit).
GN   LAMA3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Keratinocytes;
RX   MEDLINE=94357926; PubMed=8077230;
RA   Ryan M.C., Tizard R., Vandevanter D.R., Carter W.G.;
RT   "Cloning of the LamA3 gene encoding the alpha 3 chain of the adhesive
RT   ligand epiligrin. Expression in wound repair.";
RL   J. Biol. Chem. 269:22779-22787(1994).
RN   [2]
RP   SEQUENCE OF 1-1331 FROM N.A. (ISOFORMS A AND B).
RX   MEDLINE=96163880; PubMed=8586427;
RA   Vidal F., Baudoin C., Miquel C., Galliano M.-F., Christiano A.M.,
RA   Uitto J., Ortonne J.-P., Meneguzzi G.;
RT   "Cloning of the laminin alpha 3 chain gene (LAMA3) and identification
RT   of a homozygous deletion in a patient with Herlitz junctional
RT   epidermolysis bullosa.";
RL   Genomics 30:273-280(1995).
CC   -!- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
CC       IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
CC       CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
CC       WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC   -!- FUNCTION: LAMININ-5 IS THOUGHT TO BE INVOLVED IN (1) CELL ADHESION
CC       VIA INTEGRIN ALPHA-3/BETA-1 IN FOCAL ADHESION AND INTEGRIN ALPHA-
CC       6/BETA-4 IN HEMIDESMOSOMES, (2) SIGNAL TRANSDUCTION VIA TYROSINE
CC       PHOSPHORYLATION OF PP125-FAK AND P80, (3) DIFFERENTIATION OF
CC       KERATINOCYTES.
CC   -!- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
CC       DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
CC       TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
CC       COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC       THE ALPHA-3 CHAIN IS A SUBUNIT OF LAMININ-5 (EPILIGRIN/KALININ/
CC       NICEIN), AND POSSIBLY ALSO A COMPONENT OF LAMININ-6 (K-LAMININ)
CC       AND LAMININ-7 (KS-LAMININ).
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
CC       MEMBRANES (MAJOR COMPONENT).
CC   -!- ALTERNATIVE PRODUCTS: THE TWO ISOFORMS A AND B DIFFER IN THEIR N-
CC       TERMINUS. THE SEQUENCE SHOWN HERE IS THAT OF THE SMALLER VARIANT
CC       A.
CC   -!- TISSUE SPECIFICITY: SKIN; RESPIRATORY, URINARY, AND DIGESTIVE
CC       EPITHELIA AND IN OTHER SPECIALIZED TISSUES WITH PROMINENT
CC       SECRETORY OR PROTECTIVE FUNCTIONS. EPITHELIAL BASEMENT MEMBRANE,
CC       AND EPITHELIAL CELL TONGUE THAT MIGRATES INTO A WOUND BED. A
CC       DIFFERENTIAL AND FOCAL EXPRESSION OF THE ALPHA-3 CHAIN IS OBSERVED
CC       IN THE CNS.
CC   -!- INDUCTION: LAMININ-5 IS UP-REGULATED IN WOUND SITES OF HUMAN SKIN.
```

```
CC   -!- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC       WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC   -!- DOMAIN: DOMAIN G IS GLOBULAR.
CC   -!- DISEASE: DEFECTS IN LAMA3 ARE A CAUSE OF JUNCTIONAL EPIDERMOLYSIS
CC       BULLOSA (JEB) GRAVIS, ALSO KNOWN AS HERLITZ TYPE JEB. IT IS A
CC       BLISTERING DISORDER IN SKIN THAT IS CHARACTERIZED BY A SEPARATION
CC       OF BASAL CELLS FROM THE BASEMENT MEMBRANE DUE TO A DECREASED
CC       NUMBER OF HEMIDESMOSOMES. LAMININ-5 IS MISSING FROM THE BASEMENT
CC       MEMBRANE OF PATIENTS WITH THE GRAVIS FORM OF EPIDERMOLYSIS
CC       BULLOSA.
CC   -!- SIMILARITY: CONTAINS 2.5 LAMININ EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L34155; AAA59483.1; -.
DR   EMBL; X85107; CAA59428.1; -.
DR   EMBL; X85108; CAA59429.1; -.
DR   HSSP; P02468; 1TLE.
DR   MIM; 600805; -.
DR   MIM; 226700; -.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR002049; Laminin_EGF.
DR   InterPro; IPR001791; Laminin_G.
DR   InterPro; IPR003129; TSPN.
DR   Pfam; PF00053; laminin_EGF; 2.
DR   Pfam; PF00054; laminin_G; 2.
DR   SMART; SM00180; EGF_Lam; 2.
DR   SMART; SM00282; LamG; 5.
DR   SMART; SM00210; TSPN; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01248; LAMININ_TYPE_EGF; 2.
DR   PROSITE; PS50025; LAM_G_DOMAIN; 5.
KW   Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW   Laminin EGF-like domain; Cell adhesion; Repeat; Signal;
KW   Alternative splicing; Epidermolysis bullosa.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21   1713       LAMININ ALPHA-3 CHAIN.
FT   DOMAIN       46    201       DOMAIN III A.
FT   DOMAIN       67    113       LAMININ EGF-LIKE 1.
FT   DOMAIN      114    166       LAMININ EGF-LIKE 2.
FT   DOMAIN      167    185       LAMININ EGF-LIKE 3 (INCOMPLETE).
FT   DOMAIN      186    769       DOMAIN II AND I.
FT   DOMAIN      770    971       LAMININ G-LIKE 1.
FT   DOMAIN      978   1140       LAMININ G-LIKE 2.
FT   DOMAIN     1147   1307       LAMININ G-LIKE 3.
FT   DOMAIN     1366   1530       LAMININ G-LIKE 4.
FT   DOMAIN     1537   1710       LAMININ G-LIKE 5.
FT   DOMAIN      231    327       COILED COIL (POTENTIAL).
FT   DOMAIN      396    548       COILED COIL (POTENTIAL).
FT   DOMAIN      594    621       COILED COIL (POTENTIAL).
FT   DOMAIN      702    765       COILED COIL (POTENTIAL).
FT   DOMAIN     1686   1713       COILED COIL (POTENTIAL).
FT   DISULFID     67     76       BY SIMILARITY.
FT   DISULFID     69     83       BY SIMILARITY.
FT   DISULFID     86     95       BY SIMILARITY.
FT   DISULFID     98    111       BY SIMILARITY.
FT   DISULFID    114    126       BY SIMILARITY.
FT   DISULFID    116    135       BY SIMILARITY.
FT   DISULFID    137    146       BY SIMILARITY.
FT   DISULFID    149    164       BY SIMILARITY.
FT   DISULFID    202    202       INTERCHAIN (PROBABLE).
FT   DISULFID    205    205       INTERCHAIN (PROBABLE).
FT   SITE        658    660       CELL ATTACHMENT SITE (POTENTIAL).
FT   CARBOHYD    542    542       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    645    645       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    745    745       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    882    882       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    964    964       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1108   1108       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1131   1131       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1325   1325       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1477   1477       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1667   1667       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC      1     45       MGWLWIFGAALGQCLGYSSQQQRVPFLQPPGQSQLQASYVE
FT                                FRPS -> KVSSYGGYLTYQAKSFALPGDMVLLEKKPDVQL
FT                                TGQHMSIIYEETNTPRPDRLHHGRVHVVEGNFRHASSRAPV
FT                                SREELMTVLSRLADVRIQGLYFTETQRLTLSEVGLEEASDT
FT                                GSGRIALAVEICACPPAYAGDSC (IN ISOFORM B).
FT   CONFLICT      5      5       W -> R (IN REF. 2).
FT   CONFLICT    123    125       ATG -> GMC (IN REF. 2).
FT   CONFLICT    481    481       M -> K (IN REF. 2).
FT   CONFLICT    754    754       R -> L (IN REF. 2).
FT   CONFLICT    969    969       E -> Q (IN REF. 2).
FT   CONFLICT   1052   1052       D -> A (IN REF. 2).
FT   CONFLICT   1184   1184       G -> A (IN REF. 2).
SQ   SEQUENCE   1713 AA;  189304 MW;  45EA9BE1017B60D3 CRC64;


  Query Match             19.0%;  Score 68.5;  DB 1;  Length 1713;
  Best Local Similarity   35.5%;  Pred. No. 13;
  Matches   22;  Conservative   12;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       8 AVDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPEKYLFQFHNLNSNLDRIF 64
           |||  :|      :  || :| |  ||: | ||| :|:| :| |       |:|| |::
Db     552 AVDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLL 607

Qy      65 DK 66
           ::
Db     608 NE 609


Search completed: October 10, 2002, 02:09:53
Job time: 1351 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:48:45 ; Search time 69.63 Seconds
(without alignments)
96.600 Million cell updates/sec

Title: US-09-289-346A-8
Perfect score: 361
Sequence: 1 TLVWGEAAVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	350	97.0	352	1	QQCVL1	AL1 protein - toma
2	250	69.3	361	1	QQCVPT	AL1 protein - pota
3	237	65.7	358	2	S07594	hypothetical prote
4	230	63.7	362	1	JQ1887	AL1 protein - toma
5	223	61.8	359	2	S39211	gene C1 protein -
6	219	60.7	349	2	JQ2300	replicase - pepper
7	219	60.7	349	2	S31875	AL1 protein - pepp
8	216	59.8	359	2	S22593	hypothetical prote
9	213	59.0	351	2	JQ2327	AL1 protein - Indi
10	213	59.0	355	1	QQCVW1	AV1 protein - abut
11	212	58.7	357	1	QQCVC1	AL1 protein - toma
12	210	58.2	358	1	JQ1870	AL1 protein - toma
13	210	58.2	359	2	S39235	gene C1 protein -
14	210	58.2	385	2	S28360	AL1 protein - beet
15	203	56.2	360	2	S59885	replication-associ
16	142	39.3	131	2	S45059	AC1 protein (clone
17	117	32.4	347	1	QQCVS1	AL1 protein - squa
18	68.5	19.0	1713	2	A55347	adhesive ligand ep
19	68	18.8	587	2	JC1419	Fc gamma (IgG) rec
20	67	18.6	180	2	D84082	hypothetical prote
21	66.5	18.4	201	2	A81380	probable flagellar
22	64.5	17.9	299	2	B71967	probable peptidyl-
23	63.5	17.6	447	2	T12544	hypothetical prote
24	63.5	17.6	1610	2	A46227	voltage-dependent
25	63.5	17.6	1646	2	JH0422	voltage-dependent
26	63.5	17.6	2161	2	JH0564	calcium channel al
27	63.5	17.6	2181	2	A38198	calcium channel al
28	63.5	17.6	2203	2	T42742	voltage-dependent
29	62	17.2	154	2	AD3475	nitrogen regulator
30	62	17.2	181	2	G97976	conserved hypothet
31	62	17.2	295	2	D42452	C1 protein - tobac
32	62	17.2	714	2	C95382	probable ferrichro
33	61.5	17.0	481	2	A70091	probable phosphoes
34	61	16.9	1502	1	RGBYH1	CYC1/CYP3 transcri
35	60	16.6	432	2	D75348	gamma-glutamyl pho
36	60	16.6	447	2	S52437	CDP-diacylglycerol
37	59.5	16.5	136	2	T22240	hypothetical prote
38	59.5	16.5	269	2	JC7700	38K ribosome-assoc
39	59.5	16.5	299	2	G64541	cell binding facto
40	59.5	16.5	388	2	C69196	hypothetical prote
41	59.5	16.5	642	2	D90558	topoisomerase iv s
42	59.5	16.5	1033	2	E97700	hypothetical prote
43	59.5	16.5	2137	2	T05244	hypothetical prote
44	59	16.3	316	2	C82085	conserved hypothet
45	59	16.3	397	2	B71078	probable NADH oxid

ALIGNMENTS

RESULT 1
QQCVL1
AL1 protein - tomato golden mosaic virus
C;Species: tomato golden mosaic virus
A;Note: host Nicotiana sp. (tobacco)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C;Accession: A04170
R;Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.
EMBO J. 3, 2197-2205, 1984
A;Title: Complete nucleotide sequence of the infectious cloned DNA components of toma
A;Reference number: A04163
A;Accession: A04170
A;Molecule type: DNA
A;Residues: 1-352 <HAM>
C;Comment: The genome consists of two circular, single-stranded DNA components, DNA A
.
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 97.0%; Score 350; DB 1; Length 352;
Best Local Similarity 97.1%; Pred. No. 1.6e-31;
Matches 68; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
QQCVPT
AL1 protein - potato yellow mosaic virus (isolate Venezuela)
C;Species: potato yellow mosaic virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C;Accession: JU0364
R;Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.
J. Gen. Virol. 72, 1515-1520, 1991
A;Title: The nucleotide sequence of the infectious cloned DNA components of potato ye
A;Reference number: JU0362; MUID:91311403
A;Accession: JU0364
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-361 <COU>
A;Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             69.3%;  Score 250;  DB 1;  Length 361;
  Best Local Similarity   68.1%;  Pred. No. 2.7e-20;
  Matches   47;  Conservative   11;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||   :||||||||| || |||||||||: :|| |::||:||:|||:|||:|||: ||
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 DRIFDKTPE 69
           |||| | ||
Db     170 DRIFMKAPE 178


RESULT   3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C;Species: cassava latent virus
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C;Accession: S07594
R;Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A;Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A;Reference number: S07590; MUID:90174930
A;Accession: S07594
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <MOR>
A;Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C;Genetics:
A;Map position: segment DNA1
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             65.7%;  Score 237;  DB 2;  Length 358;
  Best Local Similarity   61.4%;  Pred. No. 7.8e-19;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:  :|||||||| |::||| |:||: || ||| :||| :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   4
JQ1887
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1887
R;Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A;Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A;Reference number: JQ1885; MUID:93139778
A;Accession: JQ1887
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-362 <DRY>
A;Cross-references: GB:S53251
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             63.7%;  Score 230;  DB 1;  Length 362;
  Best Local Similarity   54.1%;  Pred. No. 4.8e-18;
  Matches   46;  Conservative    9;  Mismatches   14;  Indels   16;  Gaps    1;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || |||  :|||||||| |::||| |:|||  || ||| ::||  |: |: |||||||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   5
S39211
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C;Accession: S39211
R;Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A;Description: High similarity among the tomato yellow leaf curl virus isolates from
A;Reference number: S39209
A;Accession: S39211
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <NOR>
A;Cross-references: EMBL:Z25751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             61.8%;  Score 223;  DB 2;  Length 359;
  Best Local Similarity   59.4%;  Pred. No. 2.9e-17;
  Matches   41;  Conservative    9;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | ||   :|||||||| ||:||| |:|:|| || ||| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   6
JQ2300
replicase - pepper huasteco virus (component A)
N;Alternate names: ORF AL1 protein
C;Species: pepper huasteco virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C;Accession: JQ2300
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A;Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A;Reference number: JQ2299; MUID:94015007
A;Accession: JQ2300
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.7%;  Score 219;  DB 2;  Length 349;
  Best Local Similarity   58.6%;  Pred. No. 7.8e-17;
  Matches   41;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||  :|||||||| |::||  |:|||::| |||||||:|: |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   7
S31875
AL1 protein - pepper rizado amarillo virus
C;Species: pepper rizado amarillo virus
```

```
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Sep-1999
C;Accession: S31875
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera-Bu
submitted to the EMBL Data Library, February 1993
A;Description: Complete nucleotide sequence of pepper huasteco virus: analysis and compa
A;Reference number: S31872
A;Accession: S31875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
A;Note: the source is designated as pepper huasteco virus
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.7%;  Score 219;  DB 2;  Length 349;
  Best Local Similarity   58.6%;  Pred. No. 7.8e-17;
  Matches   41;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||  :||||||||| |::|| |:|||::| |||||||:|: |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   8
S22593
hypothetical protein C4 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Sep-1999
C;Accession: S22593
R;Kheyr-Pour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.
Nucleic Acids Res. 19, 6763-6769, 1991
A;Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monoparti
A;Reference number: S22588; MUID:92107660
A;Accession: S22593
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-359 <KHE>
A;Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.8%;  Score 216;  DB 2;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 1.8e-16;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | ||   :||||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   9
JQ2327
AL1 protein - Indian cassava mosaic virus
N;Alternate names: replication-associated protein
C;Species: Indian cassava mosaic virus
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C;Accession: JQ2327; S35883
R;Hong, Y.G.; Robinson, D.J.; Harrison, B.D.
J. Gen. Virol. 74, 2437-2443, 1993
A;Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-tran
A;Reference number: JQ2326; MUID:94065670
A;Accession: JQ2327
A;Molecule type: DNA
A;Residues: 1-351 <HON>
A;Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.0%;  Score 213;  DB 2;  Length 351;
  Best Local Similarity   61.2%;  Pred. No. 3.7e-16;
  Matches   41;  Conservative    9;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       4 WGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           ||   :||||||||| |::||| | ||||: || |||:|:||  |  ||  ||:::||||||
Db     113 WGTFQIDGRSARGGQQSANDAYAAALNSGSKSEALKILRELAPRDYLRDFHHISSNLDRI 172

Qy      64 FDKTPEP 70
           | | | |
Db     173 FTKPPPP 179


RESULT  10
QQCVW1
AV1 protein - abutilon mosaic virus (isolate West India)
C;Species: abutilon mosaic virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Apr-1994
C;Accession: A36214
R;Frischmuth, T.; Zimmat, G.; Jeske, H.
Virology 178, 461-468, 1990
A;Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as
A;Reference number: A36214; MUID:91020984
A;Accession: A36214
A;Molecule type: DNA
A;Residues: 1-355 <FRI>
A;Cross-references: EMBL:X15983
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.0%;  Score 213;  DB 1;  Length 355;
  Best Local Similarity   57.1%;  Pred. No. 3.8e-16;
  Matches   40;  Conservative   12;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |  |||  :||||||||| ||:||: |:||||   : || |::|: |: |: | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
QQCVC1
AL1 protein - tomato yellow leaf curl virus
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C;Accession: D40779
R;Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.
Virology 185, 151-161, 1991
A;Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin
A;Reference number: A40779; MUID:92024070
A;Accession: D40779
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-357 <NAV>
A;Cross-references: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             58.7%;  Score 212;  DB 1;  Length 357;
  Best Local Similarity   65.6%;  Pred. No. 4.9e-16;
  Matches   40;  Conservative   11;  Mismatches   10;  Indels    0;  Gaps    0;
```

```
Qy      4 WGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
          :| : :|||||||| |::||| |||||: || ||| |::|| |: |: |||||:||||||
Db    111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy     64 F 64
          |
Db    171 F 171
```

RESULT 12
JQ1870
AL1 protein - tomato mottle virus (isolate Florida)
C;Species: tomato mottle virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1870
R;Abouzid, A.M.; Polston, J.E.; Hiebert, E.
J. Gen. Virol. 73, 3225-3229, 1992
A;Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated from
A;Reference number: JQ1869; MUID:93107858
A;Accession: JQ1870
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <ABO>
A;Cross-references: GB:L14460
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             58.2%;  Score 210;  DB 1;  Length 358;
  Best Local Similarity   54.3%;  Pred. No. 8.2e-16;
  Matches   38;  Conservative   16;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:  :|||||||| |::||: |:|||| | : || ::||: |: :: | ||: |||
Db    107 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 166

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    167 ERIFAKAPEP 176
```

RESULT 13
S39235
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Sep-1997
C;Accession: S39235
R;Crespi, S.; Noris, E.; Vaira, A.; Bosco, D.; Accotto, G.
submitted to the EMBL Data Library, December 1993
A;Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.
A;Reference number: S39233
A;Accession: S39235
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <CRE>
A;Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             58.2%;  Score 210;  DB 2;  Length 359;
  Best Local Similarity   55.1%;  Pred. No. 8.3e-16;
  Matches   38;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      2 LVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | ||   :|||||||| ||:||| |:|:|| || ||| :|:: |  |:  |||::||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINARSKSEALDVIKQLAPRDYVLHFHNISSNLD 170

Qy     62 RIFDKTPEP 70
          ::|  | |
Db    171 KVFQVPPAP 179
```

RESULT 14
S28360
AL1 protein - beet curly top virus
C;Species: beet curly top virus
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
C;Accession: S28360
R;Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
EMBO J. 5, 1761-1767, 1986
A;Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly
A;Reference number: S28360
A;Accession: S28360
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-385 <STA>
A;Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             58.2%;  Score 210;  DB 2;  Length 385;
  Best Local Similarity   54.3%;  Pred. No. 8.9e-16;
  Matches   38;  Conservative   16;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||  :|||||||| ||:||: |:|||:| ::||||::|: |: |  | ||| :|
Db    137 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 196

Qy     61 DRIFDKTPEP 70
          :|| : |:|
Db    197 QKIFQRPPDP 206
```

RESULT 15
S59885
replication-associated protein C1 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
C;Accession: S59885
R;Hong, Y.; Harrison, B.D.
submitted to the EMBL Data Library, February 1995
A;Description: Nucleotide sequences from tomato leaf curl viruses from different coun
d geminiviruses.
A;Reference number: S58346
A;Accession: S59885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <HON>
A;Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA88229.1; PID:g974211
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             56.2%;  Score 203;  DB 2;  Length 360;
  Best Local Similarity   59.1%;  Pred. No. 5e-15;
  Matches   39;  Conservative   12;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      4 WGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
          :|   :|||||||| |::||| |||:|: ||  || |:||| |: :: ||||||:|||||
Db    113 FGVFQIDGRSARGGQQSANDAYAEAINSGSKAAALDILREKAPKDFVLQFHNLNANLDRI 172

Qy     64 FDKTPE 69
          |  : |
Db    173 FTPSAE 178
```

Search completed: October 10, 2002, 01:48:45
Job time: 563 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:47:19 ; Search time 133.15 Seconds
(without alignments)
58.394 Million cell updates/sec

Title: US-09-289-346A-8
Perfect score: 361
Sequence: 1 TLVWGEAAVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	361	100.0	70	21	AAB18684	Mutant peptide der
2	350	97.0	70	21	AAB18677	Peptide fragment f
3	350	97.0	356	21	AAB18687	Amino acid sequenc
4	342	94.7	70	21	AAB18685	Mutant peptide der
5	339	93.9	70	21	AAB18688	Mutant peptide der
6	338	93.6	70	21	AAB18692	Mutant peptide der
7	336	93.1	70	21	AAB18690	Mutant peptide der
8	335	92.8	70	21	AAB18678	Mutant peptide der
9	335	92.8	70	21	AAB18686	Mutant peptide der
10	334	92.5	70	21	AAB18689	Mutant peptide der
11	332	92.0	70	21	AAB18680	Mutant peptide der
12	332	92.0	70	21	AAB18691	Mutant peptide der
13	329	91.1	70	21	AAB18681	Mutant peptide der
14	328	90.9	70	21	AAB18683	Mutant peptide der
15	326	90.3	70	21	AAB18682	Mutant peptide der
16	318	88.1	70	21	AAB18679	Mutant peptide der
17	216	59.8	353	18	AAW34338	Bean golden mosaic
18	216	59.8	353	18	AAW34332	Bean golden mosaic
19	216	59.8	353	18	AAW34333	Bean golden mosaic
20	216	59.8	353	18	AAW34334	Bean golden mosaic
21	216	59.8	353	18	AAW34335	Bean golden mosaic
22	216	59.8	359	17	AAR88870	Sardinian tomato y
23	216	59.8	359	17	AAR88871	Sardinian tomato y
24	216	59.8	359	17	AAR88872	Sardinian tomato y
25	214	59.3	353	8	AAP70407	ORF 4 gene product
26	214	59.3	361	18	AAW34336	Tomato mottle viru
27	214	59.3	361	18	AAW34324	Tomato mottle viru
28	214	59.3	361	18	AAW34325	Tomato mottle viru
29	214	59.3	361	18	AAW34326	Tomato mottle viru
30	212	58.7	357	18	AAW34329	Tomato yellow leaf
31	212	58.7	357	18	AAW34330	Tomato yellow leaf
32	212	58.7	357	18	AAW34331	Tomato yellow leaf
33	207.5	57.5	361	8	AAP70562	Product of ORF 4 f
34	204	56.5	357	18	AAW34337	Tomato yellow leaf
35	202	56.0	362	19	AAW56495	Tobacco leaf curli
36	71.5	19.8	945	22	ABB65231	Drosophila melanog
37	68.5	19.0	1693	21	AAB48457	Human laminin 5 po
38	68.5	19.0	1693	21	AAB48459	Human laminin 5 po
39	68.5	19.0	1713	16	AAR70148	Deduced sequence o
40	68.5	19.0	1713	21	AAB48458	Human laminin 5 po
41	68.5	19.0	1724	21	AAB48456	Human laminin 5 po
42	66.5	18.4	512	19	AAW68473	HIV-1 strain YBF30
43	66.5	18.4	665	22	ABB70991	Drosophila melanog
44	65	18.0	292	21	AAY92317	Geminivirus Rep C1
45	65	18.0	335	21	AAY92318	Mastrevirus Rep in

ALIGNMENTS

RESULT 1
AAB18684
ID AAB18684 standard; peptide; 70 AA.
XX
AC AAB18684;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 7
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 8
FT /note= "wild type residue replaced with Ala"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;

```
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             100.0%;  Score 361;  DB 21;  Length 70;
  Best Local Similarity   100.0%;  Pred. No. 6.2e-38;
  Matches   70;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 tlvwgeaavdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   2
AAB18677
ID   AAB18677 standard; peptide; 70 AA.
XX
AC   AAB18677;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Peptide fragment from Rep protein of TGMV (amino acids 110-179).
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Tomato golden mosaic virus.
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 18; 73pp; English.
XX
CC   The present sequence is derived from a geminivirus replication (Rep)
CC   protein, which is also known as AL1. AL1 binds double-stranded DNA,
CC   catalyses cleavage and ligation of single-stranded DNA, and interacts
CC   with other viral and host proteins. Mutants of the AL1 protein are used
CC   to produce transgenic plants. The mutation in AL1 is present in a
CC   ribosome binding region, and expression of mutant AL1 protein imparts
CC   increased resistance to geminivirus infection in the plant. Mutant AL1
CC   proteins are useful for producing plants having increased resistance or
CC   reduced sensitivity to a geminivirus such as tomato golden mosaic virus,
CC   tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl
CC   virus, African cassava mosaic virus, Indian cassava mosaic virus, potato
CC   yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,
CC   squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or
CC   beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             97.0%;  Score 350;  DB 21;  Length 70;
  Best Local Similarity   97.1%;  Pred. No. 1.5e-36;
  Matches   68;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   3
AAB18687
ID   AAB18687 standard; peptide; 356 AA.
XX
AC   AAB18687;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Amino acid sequence of a geminivirus replication protein of TGMV.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 354
FT                   /note= "unspecified amino acid"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
```

```
PT   -
XX
PS   Disclosure; Page 47-48; 73pp; English.
XX
CC   The present sequence represents a geminivirus replication (Rep)
CC   protein, which is also known as AL1. AL1 binds double-stranded DNA,
CC   catalyses cleavage and ligation of single-stranded DNA, and interacts
CC   with other viral and host proteins. Mutants of the AL1 protein are used
CC   to produce transgenic plants. The mutation in AL1 is present in a
CC   ribosome binding region, and expression of mutant AL1 protein imparts
CC   increased resistance to geminivirus infection in the plant. Mutant AL1
CC   proteins are useful for producing plants having increased resistance or
CC   reduced sensitivity to a geminivirus such as tomato golden mosaic virus,
CC   tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl
CC   virus, African cassava mosaic virus, Indian cassava mosaic virus, potato
CC   yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,
CC   squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or
CC   beet curly top virus.
XX
SQ   Sequence   356 AA;


  Query Match             97.0%;  Score 350;  DB 21;  Length 356;
  Best Local Similarity   97.1%;  Pred. No. 1.2e-35;
  Matches   68;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     110 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 169

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     170 drifdktpep 179


RESULT   4
AAB18685
ID   AAB18685 standard; peptide; 70 AA.
XX
AC   AAB18685;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 10
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 46; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             94.7%;  Score 342;  DB 21;  Length 70;
  Best Local Similarity   95.7%;  Pred. No. 1.5e-35;
  Matches   67;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||  | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 tlvwgefqvagrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT   5
AAB18688
ID   AAB18688 standard; peptide; 70 AA.
XX
AC   AAB18688;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 19
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 20
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
```

```
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 48; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             93.9%;  Score 339;  DB 21;  Length 70;
  Best Local Similarity   94.3%;  Pred. No. 3.7e-35;
  Matches   66;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          ||||||  ||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db      1 tlvwgefqvdgrsarggcaasndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   6
AAB18692
ID   AAB18692 standard; peptide; 70 AA.
XX
AC   AAB18692;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 66
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 69
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 50; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             93.6%;  Score 338;  DB 21;  Length 70;
  Best Local Similarity   94.3%;  Pred. No. 4.9e-35;
  Matches   66;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          |||||  || |
Db     61 drifdatpap 70


RESULT   7
AAB18690
ID   AAB18690 standard; peptide; 70 AA.
XX
AC   AAB18690;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 27
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 30
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
```

```
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             93.1%;  Score 336;  DB 21;  Length 70;
  Best Local Similarity   94.3%;  Pred. No. 8.7e-35;
  Matches   66;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          ||||||  |||||||||||||||||| || ||||||||||||||||||||||||||||||
Db      1 tlvwgefqvdgrsarggcqtsndaaaaalaasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   8
AAB18678
ID   AAB18678 standard; peptide; 70 AA.
XX
AC   AAB18678;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 12
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 13
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 15
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 42-43; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             92.8%;  Score 335;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 1.2e-34;
  Matches   65;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          ||||||  ||| :| ||||||||||||||||||||||||||||||||||||||||||||||
Db      1 tlvwgefqvdgaaaaggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   9
AAB18686
ID   AAB18686 standard; peptide; 70 AA.
XX
AC   AAB18686;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 24
FT                   /note= "wild type residue replaced with Leu"
FT   Misc-difference 25
FT                   /note= "wild type residue replaced with Leu"
FT   Misc-difference 26
FT                   /note= "wild type residue replaced with Leu"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
```

```
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 46; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             92.8%;  Score 335;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 1.2e-34;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          ||||||  ||||||||||||||   ||||||||||||||||||||||||||||||||||
Db      1 tlvwgefqvdgrsarggcqtsndlllealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT  10
AAB18689
ID   AAB18689 standard; peptide; 70 AA.
XX
AC   AAB18689;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 22
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 23
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 48-49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             92.5%;  Score 334;  DB 21;  Length 70;
  Best Local Similarity   94.3%;  Pred. No. 1.6e-34;
  Matches   66;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          ||||||  ||||||||||||  ||||||||||||||||||||||||||||||||||||
Db      1 tlvwgefqvdgrsarggcqtsaaaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT  11
AAB18680
ID   AAB18680 standard; peptide; 70 AA.
XX
AC   AAB18680;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 42
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 43
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 44
```

```
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 43-44; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             92.0%;  Score 332;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 2.8e-34;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||  ||||||||||||||||||||||||||||||   ||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiiaaaipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  12
AAB18691
ID   AAB18691 standard; peptide; 70 AA.
XX
AC   AAB18691;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 34
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 35
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 36
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             92.0%;  Score 332;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 2.8e-34;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||  ||||||||||||||||||||||||||   ||||||||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnassaaaalqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  13
AAB18681
ID   AAB18681 standard; peptide; 70 AA.
XX
AC   AAB18681;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
```

```
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 47
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 48
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 49
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 44; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             91.1%;  Score 329;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 6.7e-34;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||  ||||||||||||||||||||||||||||||||||||   ||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipaaalfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  14
AAB18683
ID   AAB18683 standard; peptide; 70 AA.
XX
AC   AAB18683;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 59
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 61
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 62
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             90.9%;  Score 328;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 8.9e-34;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||  |||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsal 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 aaifdktpep 70


RESULT  15
AAB18682
ID   AAB18682 standard; peptide; 70 AA.
XX
AC   AAB18682;
XX
```

DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 52
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 54
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 55
FT /note= "wild type residue replaced with Ala"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DR WPI; 2000-618851/59.
XX
PT Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT -
XX
PS Claim 53; Page 44-45; 73pp; English.
XX
CC The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
SQ Sequence 70 AA;

Query Match 90.3%; Score 326; DB 21; Length 70;
Best Local Similarity 92.9%; Pred. No. 1.6e-33;
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||  ||||||||||||||||||||||||||||||||||||||| |  |||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfafaalnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70
```

Search completed: October 10, 2002, 01:47:19

Job time: 528 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:02:00 ; Search time 118.38 Seconds
(without alignments)
102.295 Million cell updates/sec

Title: US-09-289-346A-9
Perfect score: 362
Sequence: 1 TLVWGEFQVAGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	307	84.8	351	12	Q91R10	Q91r10 tomato seve
2	305	84.3	352	12	Q9E000	Q9e000 tomato rugo
3	288	79.6	226	12	O09727	O09727 leonurus mo
4	288	79.6	226	12	Q9WHF6	Q9whf6 tomato mild
5	286	79.0	361	12	Q67574	Q67574 bean golden
6	284	78.5	225	12	Q9QDB1	Q9qdb1 cowpea gold
7	284	78.5	314	12	Q9ELT8	Q9elt8 sweet potat
8	284	78.5	364	12	Q9QS55	Q9qs55 sweet potat
9	281	77.6	185	12	Q98693	Q98693 sida golden
10	273	75.4	149	12	P88975	P88975 macroptiliu
11	273	75.4	233	12	Q9YLA4	Q9yla4 macroptiliu
12	267	73.8	234	12	O39180	O39180 geminivirid
13	264	72.9	190	12	Q9Z089	Q9z089 tobacco lea
14	264	72.9	190	12	Q9Z084	Q9z084 tobacco lea
15	263	72.7	190	12	Q9W827	Q9w827 tobacco lea
16	263	72.7	208	12	Q9Z0C4	Q9z0c4 tobacco lea
17	263	72.7	359	12	Q91M88	Q91m88 tobacco lea
18	261	72.1	208	12	Q9Z0C0	Q9z0c0 tobacco lea
19	261	72.1	208	12	Q9Z0B8	Q9z0b8 tobacco lea
20	259	71.5	208	12	Q9Z0B6	Q9z0b6 tobacco lea
21	258	71.3	203	12	Q9Z083	Q9z083 tobacco lea
22	258	71.3	363	12	O73577	O73577 cotton leaf
23	253	69.9	359	12	Q91B86	Q91b86 ageratum ye
24	253	69.9	360	12	Q9DXI0	Q9dxi0 ageratum ye
25	251	69.3	208	12	Q9Z0A0	Q9z0a0 tobacco lea
26	251	69.3	363	12	O72705	O72705 cotton leaf
27	251	69.3	363	12	O72719	O72719 cotton leaf
28	250	69.1	190	12	Q9Z086	Q9z086 tobacco lea
29	250	69.1	208	12	Q9Z0C6	Q9z0c6 tobacco lea
30	249	68.8	349	12	Q88888	Q88888 tomato pseu
31	249	68.8	362	12	O56816	O56816 chayote mos
32	248	68.5	359	12	Q9YZV4	Q9yzv4 tomato yell
33	248	68.5	359	12	Q9YZV2	Q9yzv2 tomato yell
34	248	68.5	359	12	Q9YUX7	Q9yux7 tomato yell
35	248	68.5	359	12	Q9YL27	Q9yl27 tomato yell
36	247	68.2	348	12	Q911W5	Q911w5 macroptiliu
37	247	68.2	353	12	O72692	O72692 beet curly
38	247	68.2	354	12	Q65438	Q65438 beet curly
39	247	68.2	359	12	Q88942	Q88942 tomato yell
40	246	68.0	190	12	Q9Z0A7	Q9z0a7 tobacco lea
41	246	68.0	362	12	Q9IN48	Q9in48 okra enatio
42	245	67.7	307	12	Q9IET7	Q9iet7 cotton leaf
43	245	67.7	361	12	O72723	O72723 cotton leaf
44	244	67.4	231	12	Q96620	Q96620 african tom
45	244	67.4	354	12	Q9IN42	Q9in42 south afric

ALIGNMENTS

RESULT 1
Q91R10
ID Q91R10 PRELIMINARY; PRT; 351 AA.
AC Q91R10;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE REP PROTEIN.
GN AC1.
OS Tomato severe rugose virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=158463;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MINAS GERAIS;
RA Rezende W.L., Goulart L.R., Parreira K.S., Figueiredo J.E.F.;
RT "The full-length DNA-A nucleotide sequence of a novel tomato-infecting
RT begomovirus, Tomato severe rugose virus, in Brazil.";
RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY029750; AAK50357.1; -.
SQ SEQUENCE 351 AA; 40122 MW; 87F937A4F873B6CF CRC64;

Query Match 84.8%; Score 307; DB 12; Length 351;
Best Local Similarity 81.4%; Pred. No. 4.8e-27;
Matches 57; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||: |||||||||:|||||||||| ||: ||||||||:|||:||||||||||||
Db     111 TIEWGEFQIDGRSARGGCQTANDAAAEALNAPSKDVALQIIREKLPEKFLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
             |||| : |||
Db     171 DRIFARAPEP 180
```

RESULT 2
Q9E000
ID Q9E000 PRELIMINARY; PRT; 352 AA.

```
AC   Q9E000;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN.
GN   AC1.
OS   Tomato rugose mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=134599;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Fernandes J.J., Fontes E.P.B., Brommonschenkel S.H., Carvalho M.G.,
RA   Zambolim E.M., Zerbini F.M.;
RT   "Molecular Cloning and Characterization of Tomato rugose mosaic virus
RT   (TRMV), a Begomovirus Isolated from Tomatoes at Triangulo Mineiro,
RT   Minas Gerais, Brazil.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF291705; AAG15546.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   352 AA;  40012 MW;  47CD55838E24D613 CRC64;


  Query Match             84.3%;  Score 305;  DB 12;  Length 352;
  Best Local Similarity   81.4%;  Pred. No. 8.2e-27;
  Matches   57;  Conservative    7;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||: |||||||||||:|||||||||| ||: ||||||||:|||:||||||||||
Db     111 TIEWGEFQIDGRSARGGCQTANDAAAEALNAPSKDIALQIIREKMPEKFLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| : |||
Db     171 DRIFARAPEP 180


RESULT   3
O09727
ID   O09727      PRELIMINARY;      PRT;   226 AA.
AC   O09727;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   REP.
OS   Leonurus mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=58177;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LEMV- BRAZIL 1;
RA   Faria J.C., Maxwell D.P.;
RT   "Variability in geminivirus associated with Phaseolus vulgaris in
RT   Brazil.";
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U92532; AAB51157.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     226    226
SQ   SEQUENCE   226 AA;  25617 MW;  73CDB6E766083FC5 CRC64;


  Query Match             79.6%;  Score 288;  DB 12;  Length 226;
  Best Local Similarity   80.0%;  Pred. No. 4.3e-25;
  Matches   56;  Conservative    4;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||| |||:||| || |||||||||| |  ||||||:||:||||||||||||||
Db     111 TVEWGEFQVDGRSSRGGQQTVNDAAAEALNAPDKRTALQIIKEKLPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180


RESULT   4
Q9WHF6
ID   Q9WHF6      PRELIMINARY;      PRT;   226 AA.
AC   Q9WHF6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   tomato mild mottle geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX   NCBI_TaxID=92943;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HN96-H5KW;
RA   Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA   Maxwell D.P.;
RT   "Molecular characterization and DNA-based detection methods for
RT   vegetable-infecting geminiviruses in Central America.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF131071; AAD33471.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     226    226
SQ   SEQUENCE   226 AA;  25941 MW;  2EA4116712871A23 CRC64;


  Query Match             79.6%;  Score 288;  DB 12;  Length 226;
  Best Local Similarity   75.7%;  Pred. No. 4.3e-25;
  Matches   53;  Conservative   12;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||: ||||||| ||:|||||||||||||||::||:||:|:::|||:|||:|||
Db     111 TIEWGEFQIDGRSARGGQQTANDAAAEALNASSKEEAMRIIKEKLPKEFLFQYHNLSSNL 170

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180


RESULT   5
Q67574
ID   Q67574      PRELIMINARY;      PRT;   361 AA.
AC   Q67574;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE  REPLICATIVE PROTEIN.
GN   AL1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,
RA   Maxwell D.P., Russell D.R.;
RT   "Cloning of the complete DNA genomes of four bean-infecting
RT   geminiviruses and determining their infectivity by electric discharge
RT   particle acceleration.";
RL   Phytopathology 81:980-985(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,
RA   Morales F.J., Maxwell D.P.;
RT   "Differentiation of bean-infecting geminiviruses by nucleic acid
```

```
RT   hybridization probes and aspects of bean golden mosaic in Brazil.";
RL   Plant Dis. 75:336-342(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RT   "Genetic diversity in geminiviruses causing bean golden mosaic
RT   disease: The nucleotide sequence of the infectious cloned DNA
RT   components of a Brazilian isolate of bean golden mosaic geminivirus.";
RL   Submitted (MAY-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; M88686; AAA46312.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   361 AA;  41041 MW;  0094C7ACAF06B788 CRC64;


  Query Match             79.0%;  Score 286;  DB 12;  Length 361;
  Best Local Similarity   79.1%;  Pred. No. 1.3e-24;
  Matches   53;  Conservative    9;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       4 WGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           || ||| ||||||| ||:||||:||||||||||:|||:||:|||:|||:|||:||||||
Db     113 WGHFQVDGRSARGGQQTANDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
           | | |:|
Db     173 FTKAPDP 179


RESULT   6
Q9QDB1
ID   Q9QDB1      PRELIMINARY;      PRT;   225 AA.
AC   Q9QDB1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   cowpea golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=69263;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CGMV-BR;
RA   Faria J.C.;
RT   "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT   Brazil.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF188708; AAF06318.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     225    225
SQ   SEQUENCE   225 AA;  25766 MW;  1089CB6BD8D15B5D CRC64;


  Query Match             78.5%;  Score 284;  DB 12;  Length 225;
  Best Local Similarity   79.1%;  Pred. No. 1.2e-24;
  Matches   53;  Conservative    8;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       4 WGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           || ||: ||||||| || ||||:||||||||||||:|||:||:|||:|||:|||:||||||
Db     113 WGHFQIDGRSARGGQQTINDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
           | | |||
Db     173 FKKPPEP 179


RESULT   7
Q9ELT8
ID   Q9ELT8      PRELIMINARY;      PRT;   314 AA.
AC   Q9ELT8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION ASSOCIATION PROTEIN.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF288227; AAG01006.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   314 AA;  35153 MW;  6B6220613046943F CRC64;


  Query Match             78.5%;  Score 284;  DB 12;  Length 314;
  Best Local Similarity   82.4%;  Pred. No. 1.8e-24;
  Matches   56;  Conservative    3;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||| ||||||| ||:|||||||||| ||| ||||||||:|||||||||| |||
Db     110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy      61 DRIFDKTP 68
           ||||   |
Db     170 DRIFSPPP 177


RESULT   8
Q9QS55
ID   Q9QS55      PRELIMINARY;      PRT;   364 AA.
AC   Q9QS55;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION INITIATION PROTEIN AC1.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A.;
RT   "Cloning of a DNA-A-like genomic component of sweet potato leaf curl
RT   virus : nucleotide sequence and phylogenetic relationships.";
RL   Online Publication.
DR   EMBL; AF104036; AAD47173.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   364 AA;  40680 MW;  5F79752431A09D6E CRC64;
```

```
  Query Match             78.5%;  Score 284;  DB 12;  Length 364;
  Best Local Similarity   82.4%;  Pred. No. 2.2e-24;
  Matches   56;  Conservative    3;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||| ||||||| ||:|||||||||| ||| |||||||:||||||||||| |||
Db     110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy      61 DRIFDKTP 68
           ||||   |
Db     170 DRIFSPPP 177


RESULT   9
Q98693
ID   Q98693      PRELIMINARY;      PRT;   185 AA.
AC   Q98693;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   AC1.
OS   sida golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51034;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA;
RA   Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RT   "Genetic Diversity among geminiviruses associated with the weed
RT   species Sida spp, Macroptilium lathyroides, and Wissadula amplissima
RT   from Jamaica.";
RL   Plant Dis. 81:1251-1258(1997).
DR   EMBL; U67926; AAB97865.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     185    185
SQ   SEQUENCE   185 AA;  20975 MW;  3913850A025A5EE1 CRC64;


  Query Match             77.6%;  Score 281;  DB 12;  Length 185;
  Best Local Similarity   72.9%;  Pred. No. 2.2e-24;
  Matches   51;  Conservative   12;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||: ||||||| ||:|||||||||: :||:||:|||||:||:||||:|||:||:
Db      89 TIEWGVFQIDGRSARGGQQTANDAAAEALNSGTKEDALKIIREKLPERYLFQYHNLSSNI 148

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     149 DRIFSKPPEP 158


RESULT  10
P88975
ID   P88975      PRELIMINARY;      PRT;   149 AA.
AC   P88975;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   AC1.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICAN;
RA   Roye M.E.;
RL   Thesis (1996), Biochemistry, University of the West Indies, Jamaica.
DR   EMBL; U75278; AAB36919.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     149    149
SQ   SEQUENCE   149 AA;  16785 MW;  E4CF5EED4C9CD508 CRC64;


  Query Match             75.4%;  Score 273;  DB 12;  Length 149;
  Best Local Similarity   71.4%;  Pred. No. 1.4e-23;
  Matches   50;  Conservative   12;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||: ||||||| |||||||||||: :|| |::|::||:|||:|||:|||:|||
Db      52 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 111

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     112 DRIFMKDPEP 121


RESULT  11
Q9YLA4
ID   Q9YLA4      PRELIMINARY;      PRT;   233 AA.
AC   Q9YLA4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E.;
RT   "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
RT   from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E., McLaughlin W.A., Maxwell D.P.;
RT   "Molecular characterization of two distinct geminiviruses infecting M.
RT   lathyroides from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF098940; AAD17850.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     233    233
SQ   SEQUENCE   233 AA;  26356 MW;  AA490AF4D2166A02 CRC64;


  Query Match             75.4%;  Score 273;  DB 12;  Length 233;
  Best Local Similarity   71.4%;  Pred. No. 2.3e-23;
  Matches   50;  Conservative   12;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||: ||||||| |||||||||||: :|| |::|::||:|||:|||:|||:|||
Db     110 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 169

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     170 DRIFMKDPEP 179


RESULT  12
O39180
```

```
ID   O39180      PRELIMINARY;      PRT:   234 AA.
AC   O39180;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   AC1 (FRAGMENT).
OS   Geminiviridae.
OC   Viruses; ssDNA viruses.
OX   NCBI_TaxID=10811;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOMATO STRAIN;
RA   Guzman P., Arredondo C.R., Emmatty D., Portillo R.J., Gilbertson R.L.;
RT   "Partial Characterization of Two Whitefly-Transmitted Geminiviruses
RT   Infecting Tomatoes in Venezuela.";
RL   Plant Dis. 81:312-312(1997).
DR   EMBL; AF026553; AAB82605.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     234    234
SQ   SEQUENCE   234 AA;  26486 MW;  9ED8F0697105CD19 CRC64;


  Query Match             73.8%;  Score 267;  DB 12;  Length 234;
  Best Local Similarity   70.0%;  Pred. No. 1.1e-22;
  Matches   49;  Conservative   13;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||: ||||||| ||:|||||||||: :||||::||:||:|||:|||:|||:|:| 
Db     110 TIEWGLFQIDGRSARGGQQTANDAAAEALNSGTKEEAMKIIKEKLPEKFLFQYHNLSSDL 169

Qy      61 DRIFDKTPEP 70
           |||| | | |
Db     170 DRIFMKAPNP 179


RESULT  13
Q9Z089
ID   Q9Z089      PRELIMINARY;      PRT:   190 AA.
AC   Q9Z089;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA3-1, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001315; BAA34033.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21432 MW;  AAC093D1D1610FAD CRC64;


  Query Match             72.9%;  Score 264;  DB 12;  Length 190;
  Best Local Similarity   61.2%;  Pred. No. 2e-22;
  Matches   52;  Conservative    8;  Mismatches    9;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || || ||: ||||||||| :||| ||||||||| ||| |||||:|: ::||:|||||||
Db      85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  14
Q9Z084
ID   Q9Z084      PRELIMINARY;      PRT:   190 AA.
AC   Q9Z084;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA5-2, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA5;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001318; BAA34039.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  AAC1C2943E3F01AD CRC64;


  Query Match             72.9%;  Score 264;  DB 12;  Length 190;
  Best Local Similarity   61.2%;  Pred. No. 2e-22;
  Matches   52;  Conservative    8;  Mismatches    9;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || || ||: ||||||||| :||| ||||||||| ||| |||||:|: ::||:|||||||
Db      85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  15
Q9W827
ID   Q9W827      PRELIMINARY;      PRT:   190 AA.
AC   Q9W827;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE GORA1-1, PARTIAL AND
DE   COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GORA1;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001303; BAA34010.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
```

```
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  93C3742A8EBDB7EB CRC64;


  Query Match             72.7%;  Score 263;  DB 12;  Length 190;
  Best Local Similarity   61.2%;  Pred. No. 2.6e-22;
  Matches   52;  Conservative    8;  Mismatches    9;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || || ||| ||||||||| :||| |||||:|| ||| |||||:|: ::||:|||||||
Db      85 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFSNSSFDQVPE 169


Search completed: October 10, 2002, 02:02:00
Job time: 1093 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:09:53 ; Search time 34.62 Seconds
(without alignments)
78.289 Million cell updates/sec

Title: US-09-289-346A-9
Perfect score: 362
Sequence: 1 TLVWGEFQVAGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score-distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	356	98.3	352	1	VAL1_TGMV	P03567 tomato gold
2	256	70.7	361	1	VAL1_PYMVV	P27258 potato yell
3	243	67.1	358	1	VAL1_CLVK	P14982 cassava lat
4	243	67.1	358	1	VAL1_CLVN	P14972 cassava lat
5	236	65.2	362	1	VAL1_TYLCA	P36279 tomato yell
6	229	63.3	359	1	VAL1_TYLCU	P38609 tomato yell
7	225	62.2	349	1	VAL1_PHUV	Q06923 pepper huas
8	222	61.3	359	1	VAL1_TYLCM	P27260 tomato yell
9	220	60.8	353	1	VAL1_BGMV	P05175 bean golden
10	219	60.5	355	1	VAL1_ABMVW	P21947 abutilon mo
11	216	59.7	358	1	VAL1_BCTV	P14991 beet curly
12	216	59.7	361	1	VAL1_TMOV	Q06657 tomato mott
13	207	57.2	357	1	VAL1_TYLCV	P27259 tomato yell
14	125	34.5	347	1	VAL1_SLCV	P29048 squash leaf
15	64.5	17.8	299	1	Y175_HELPJ	Q9zmq7 helicobacte
16	62.5	17.3	447	1	TBL2_HUMAN	Q9y4p3 homo sapien
17	62	17.1	295	1	VAL1_TYDVA	P31617 tobacco yel
18	61	16.9	1483	1	CYP1_YEAST	P12351 saccharomyc
19	60.5	16.7	1713	1	LMA3_HUMAN	Q16787 homo sapien
20	60	16.6	319	1	LDH2_STAAM	Q99r35 staphylococ
21	60	16.6	447	1	CDSA_DROME	P56079 d phosphati
22	59	16.3	1287	1	SKI2_YEAST	P35207 saccharomyc
23	58.5	16.2	297	1	RRPP_RABVP	P06747 rabies viru
24	58.5	16.2	299	1	Y175_HELPY	P56112 helicobacte
25	58.5	16.2	367	1	LHX4_MOUSE	P53776 mus musculu
26	58.5	16.2	589	1	YE67_HAEIN	P45221 haemophilus
27	58.5	16.2	1044	1	BUB1_SCHPO	O94751 schizosacch
28	58	16.0	207	1	IL6_MARMO	O35736 marmota mon
29	58	16.0	256	1	YKJ9_YEAST	P34247 saccharomyc
30	58	16.0	617	1	YACH_ECOLI	P36682 escherichia
31	58	16.0	874	1	SLAP_BACLI	P49052 bacillus li
32	57.5	15.9	136	1	Y452_CAEEL	O62250 caenorhabdi
33	57	15.7	355	1	CRTB_RHOSH	P54905 rhodobacter
34	57	15.7	492	1	MOT3_MOUSE	O35308 mus musculu
35	57	15.7	492	1	MOT3_RAT	O70461 rattus norv
36	56.5	15.6	428	1	GBA1_CANAL	P28868 candida alb
37	56.5	15.6	470	1	RHBA_RHIME	Q9z3r2 rhizobium m
38	56.5	15.6	511	1	HUTH_VIBCH	Q9ksq4 vibrio chol
39	56	15.5	332	1	IPAD_SHIFL	P18013 shigella fl
40	56	15.5	336	1	Y625_METJA	Q58042 methanococc
41	56	15.5	432	1	PROA_DEIRA	Q9rtd9 deinococcus
42	56	15.5	513	1	HEM0_CHICK	P18080 gallus gall
43	56	15.5	601	1	CYSJ_BUCAI	P57503 buchnera ap
44	56	15.5	795	1	LON_AQUAE	O66605 aquifex aeo
45	56	15.5	799	1	AFSK_STRCO	P54741 streptomyce

ALIGNMENTS

RESULT 1
VAL1_TGMV
ID VAL1_TGMV STANDARD; PRT; 352 AA.
AC P03567;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein.
GN AC1.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10831;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;
RT "Complete nucleotide sequence of the infectious cloned DNA components
RT of tomato golden mosaic virus: potential coding regions and regulatory
RT sequences.";
RL EMBO J. 3:2197-2205(1984).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; K02029; -; NOT_ANNOTATED_CDS.
DR PIR; A04170; QQCVL1.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR ProDom; PD000736; Gemini_AL1; 1.
KW ATP-binding.
FT NP_BIND 223 230 ATP (POTENTIAL).
SQ SEQUENCE 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;

Query Match 98.3%; Score 356; DB 1; Length 352;
Best Local Similarity 98.6%; Pred. No. 2.2e-32;
Matches 69; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
VAL1_PYMVV
ID VAL1_PYMVV STANDARD; PRT; 361 AA.

```
AC   P27258;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 protein.
OS   Potato yellow mosaic virus (isolate Venezuela).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10828;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91311403; PubMed=1856690;
RA   Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT   "The nucleotide sequence of the infectious cloned DNA components of
RT   potato yellow mosaic virus.";
RL   J. Gen. Virol. 72:1515-1520(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00940; BAA00782.1; -.
DR   PIR; JU0364; QQCVPT.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   361 AA;  40850 MW;  5627A33BF1264383 CRC64;


  Query Match             70.7%;  Score 256;  DB 1;  Length 361;
  Best Local Similarity   69.6%;  Pred. No. 3.1e-21;
  Matches   48;  Conservative   11;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: || ||: ||||||| || |||||||||: :|| |::|||:||:|||:|||:|||: ||
Db    110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy     61 DRIFDKTPEP 69
          |||| | ||
Db    170 DRIFMKAPE 178


RESULT   3
VAL1_CLVK
ID   VAL1_CLVK      STANDARD;      PRT;   358 AA.
AC   P14982;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (40.4 kDa protein).
GN   AC1.
OS   Cassava latent virus (strain West Kenyan 844).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10818;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Gay M.R.;
RT   "Nucleotide sequence of cassava latent virus DNA.";
RL   Nature 301:260-262(1983).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J02057; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40346 MW;  ED173E753EE92D69 CRC64;


  Query Match             67.1%;  Score 243;  DB 1;  Length 358;
  Best Local Similarity   62.9%;  Pred. No. 8.6e-20;
  Matches   44;  Conservative   13;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||: ||||||| |::||| |:|||: || ||| :||| :|: :: |||||||||
Db    109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy     61 DRIFDKTPEP 70
          |||| : | |
Db    169 DRIFQEPPAP 178


RESULT   4
VAL1_CLVN
ID   VAL1_CLVN      STANDARD;      PRT;   358 AA.
AC   P14972;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (40.4 kDa protein).
GN   AC1.
OS   Cassava latent virus (strain Nigerian).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10819;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90174930; PubMed=2308831;
RA   Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT   "Nucleotide sequence of the infectious cloned DNA components of
RT   African cassava mosaic virus (Nigerian strain).";
RL   Nucleic Acids Res. 18:197-198(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X17095; CAA34953.1; -.
DR   PIR; S07594; S07594.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40435 MW;  1DB16BB0CB2D5E2C CRC64;


  Query Match             67.1%;  Score 243;  DB 1;  Length 358;
  Best Local Similarity   62.9%;  Pred. No. 8.6e-20;
  Matches   44;  Conservative   13;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||: ||||||| |::||| |:|||: || ||| :||| :|: :: |||||||||
Db    109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168
```

```
Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   5
VAL1_TYLCA
ID   VAL1_TYLCA      STANDARD;      PRT;   362 AA.
AC   P36279;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=36447;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93139778; PubMed=8423446;
RA   Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;
RT   "Nucleotide sequence and genome organization of tomato leaf curl
RT   geminivirus.";
RL   J. Gen. Virol. 74:147-151(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
DR   PIR; JQ1887; JQ1887.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   362 AA;  41197 MW;  343E7184B4704098 CRC64;


  Query Match             65.2%;  Score 236;  DB 1;  Length 362;
  Best Local Similarity   55.3%;  Pred. No. 5.2e-19;
  Matches   47;  Conservative    9;  Mismatches   13;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || |||||: ||||||| |::||| |:||| || ||| ::|| |: |: |||||||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   6
VAL1_TYLCU
ID   VAL1_TYLCU      STANDARD;      PRT;   359 AA.
AC   P38609;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=37139;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94256836; PubMed=8198442;
RA   Noris E., Hidalgo E., Accotto G., Moriones E.;
RT   "High similarity among the tomato yellow leaf curl virus isolates
RT   from the west Mediterranean basin: the nucleotide sequence of an
RT   infectious clone from Spain.";
RL   Arch. Virol. 135:165-170(1994).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z25751; CAA81026.1; -.
DR   PIR; S39211; S39211.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  41065 MW;  2D170A51EF80A3EC CRC64;


  Query Match             63.3%;  Score 229;  DB 1;  Length 359;
  Best Local Similarity   60.9%;  Pred. No. 3.1e-18;
  Matches   42;  Conservative    9;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||: ||||||| ||:||| |:|:|| || ||| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   7
VAL1_PHUV
ID   VAL1_PHUV      STANDARD;      PRT;   349 AA.
AC   Q06923;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AL1.
OS   Pepper huasteco virus (PHV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=28349;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94015007; PubMed=8409944;
RA   Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA   Rivera-Bustamante R.F.;
RT   "Complete nucleotide sequence of pepper huasteco virus: analysis and
RT   comparison with bipartite geminiviruses.";
RL   J. Gen. Virol. 74:2225-2231(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70418; CAA49856.1; -.
DR   PIR; S31875; S31875.
DR   PIR; JQ2300; JQ2300.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (BY SIMILARITY).
SQ   SEQUENCE   349 AA;  39722 MW;  D5F4E76CD56370F4 CRC64;


  Query Match             62.2%;  Score 225;  DB 1;  Length 349;
```

```
  Best Local Similarity   60.0%;  Pred. No. 8.4e-18;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||: ||||||| |::||  |:|||::| |||||||:|: |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   8
VAL1_TYLCM
ID   VAL1_TYLCM     STANDARD;      PRT;   359 AA.
AC   P27260;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10833;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92107660; PubMed=1840676;
RA   Kheyr-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,
RA   Gronenborn B.;
RT   "Tomato yellow leaf curl virus from Sardinia is a
RT   whitefly-transmitted monopartite geminivirus.";
RL   Nucleic Acids Res. 19:6763-6769(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X61153; CAA43466.1; -.
DR   PIR; S22593; S22593.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  40733 MW;  9717B4A07C93EFA7 CRC64;


  Query Match             61.3%;  Score 222;  DB 1;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 1.9e-17;
  Matches   40;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||: ||||||| ||:||| |:|::|| || :|| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   9
VAL1_BGMV
ID   VAL1_BGMV      STANDARD;      PRT;   353 AA.
AC   P05175;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   AL1 protein (40.2 kDa protein).
GN   AC1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Howarth A.J., Caton J., Bossert M., Goodman R.M.;
RT   "Nucleotide sequence of bean golden mosaic virus and a model for gene
RT   regulation in geminiviruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M10070; AAA46318.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   353 AA;  40190 MW;  80FA779DF6029A34 CRC64;


  Query Match             60.8%;  Score 220;  DB 1;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 3.1e-17;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||| ||||||| |::||: |:|||| | | || |::|: |: |: | ||: |||
Db     110 TIEWGQFQVDGRSARGGQQSANDSYAKALNADSIESALTILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFFKVPEP 179


RESULT  10
VAL1_ABMVW
ID   VAL1_ABMVW     STANDARD;      PRT;   355 AA.
AC   P21947;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AC1.
OS   Abutilon mosaic virus (isolate West India).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10816;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91020984; PubMed=2219703;
RA   Frischmuth T., Zimmat G., Jeske H.;
RT   "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT   as well as eukaryotic features.";
RL   Virology 178:461-468(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR   PIR; A36214; QQCVW1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   355 AA;  40257 MW;  16A2CA8A63251E95 CRC64;


  Query Match             60.5%;  Score 219;  DB 1;  Length 355;
  Best Local Similarity   58.6%;  Pred. No. 4e-17;
  Matches   41;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |  |||||: ||||||| ||:||: |:|||    : || |::|: |: |: | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
VAL1_BCTV
ID   VAL1_BCTV      STANDARD;      PRT;   358 AA.
AC   P14991;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 protein (40.8 kDa protein).
OS   Beet curly top virus (BCTV).
OC   Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
OX   NCBI_TaxID=10840;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT   "The nucleotide sequence of an infectious clone of the geminivirus
RT   beet curly top virus.";
RL   EMBO J. 5:1761-1767(1986).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40889 MW;  39A45FE3C0B9C333 CRC64;


  Query Match             59.7%;  Score 216;  DB 1;  Length 358;
  Best Local Similarity   55.7%;  Pred. No. 8.7e-17;
  Matches   39;  Conservative   16;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||: ||||||| ||:||: |:||||:| ::||||::|: |: |  | ||| :|
Db     110 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 169

Qy      61 DRIFDKTPEP 70
           :|| : |:|
Db     170 QKIFQRPPDP 179


RESULT  12
VAL1_TMOV
ID   VAL1_TMOV      STANDARD;      PRT;   361 AA.
AC   Q06657;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AL1.
OS   Tomato mottle virus (isolate Florida) (TMoV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=36449;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93107858; PubMed=1469361;
RA   Abouzid A.M., Polston J.E., Hiebert E.;
RT   "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT   isolated from tomatoes in Florida.";
RL   J. Gen. Virol. 73:3225-3229(1992).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L14460; AAC32414.1; -.
DR   PIR; JQ1870; JQ1870.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (BY SIMILARITY).
SQ   SEQUENCE   361 AA;  40516 MW;  8138B65CEEAC6950 CRC64;


  Query Match             59.7%;  Score 216;  DB 1;  Length 361;
  Best Local Similarity   55.7%;  Pred. No. 8.8e-17;
  Matches   39;  Conservative   16;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||: ||||||| |::||: |:|||| | : || ::||: |: :: | ||: |||
Db     110 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  13
VAL1_TYLCV
ID   VAL1_TYLCV     STANDARD;      PRT;   357 AA.
AC   P27259;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10832;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92024070; PubMed=1926771;
RA   Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;
RT   "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RT   with a single genomic component.";
RL   Virology 185:151-161(1991).
```

```
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15656; CAA33688.1; -.
DR   PIR; D40779; QQCVC1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     219    226       ATP (POTENTIAL).
SQ   SEQUENCE   357 AA;  40678 MW;  939AB68E1AB3B2A7 CRC64;


  Query Match             57.2%;  Score 207;  DB 1;  Length 357;
  Best Local Similarity   65.6%;  Pred. No. 8.7e-16;
  Matches   40;  Conservative   10;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       4 WGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           :|  |: |||||| |::||| |||||: || ||| |::|| |: |: ||||||:|||||
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171


RESULT  14
VAL1_SLCV
ID   VAL1_SLCV      STANDARD;      PRT;   347 AA.
AC   P29048;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   AL1 protein.
OS   Squash leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10829;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91082449; PubMed=1984668;
RA   Lazarowitz S.G., Lazdins I.B.;
RT   "Infectivity and complete nucleotide sequence of the cloned genomic
RT   components of a bipartite squash leaf curl geminivirus with a broad
RT   host range phenotype.";
RL   Virology 180:58-69(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M38183; AAC32410.1; ALT_INIT.
DR   PIR; C36785; QQCVS1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     218    225       ATP (POTENTIAL).
SQ   SEQUENCE   347 AA;  39110 MW;  AFDAEBDDE122110E CRC64;


  Query Match             34.5%;  Score 125;  DB 1;  Length 347;
  Best Local Similarity   37.9%;  Pred. No. 1.2e-06;
  Matches   25;  Conservative   15;  Mismatches   22;  Indels    4;  Gaps    1;

Qy       5 GEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRIF 64
           |:::|:|    |     :|    |:|| |  ||| ||:   |: ::  :||| :|::|:|
Db     116 GQYKVSG----GSKSNKDDVYHNAVNAGSAGEALDIIKAGDPKTFIVNYHNLLANVERLF 171

Qy      65 DKTPEP 70
             | |||
Db     172 QKPPEP 177


RESULT  15
Y175_HELPJ
ID   Y175_HELPJ     STANDARD;      PRT;   299 AA.
AC   Q9ZMQ7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein JHP0161 precursor.
GN   JHP0161.
OS   Helicobacter pylori J99 (Campylobacter pylori J99).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
OX   NCBI_TaxID=85963;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99120557; PubMed=9923682;
RA   Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA   Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA   Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA   Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA   Trust T.J.;
RT   "Genomic sequence comparison of two unrelated isolates of the human
RT   gastric pathogen Helicobacter pylori.";
RL   Nature 397:176-180(1999).
CC   -!- SIMILARITY: BELONGS TO THE PPIC/PARVULIN FAMILY OF ROTAMASES.
CC       STRONG, TO C.JEJUNI CBF2.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001454; AAD05744.1; -.
DR   HSSP; Q13526; 1PIN.
DR   InterPro; IPR000297; Rotamase.
DR   Pfam; PF00639; Rotamase; 1.
DR   PROSITE; PS01096; PPIC_PPIASE_1; 1.
DR   PROSITE; PS50198; PPIC_PPIASE_2; 1.
KW   Hypothetical protein; Isomerase; Rotamase; Signal; Complete proteome.
FT   SIGNAL       1     21       POTENTIAL.
FT   CHAIN       22    299       HYPOTHETICAL PROTEIN JHP0161.
FT   DOMAIN     154    253       PPIC.
SQ   SEQUENCE   299 AA;  34040 MW;  9C037B1CD1110143 CRC64;


  Query Match             17.8%;  Score 64.5;  DB 1;  Length 299;
  Best Local Similarity   28.1%;  Pred. No. 5.3;
  Matches   18;  Conservative   14;  Mismatches   11;  Indels   21;  Gaps    2;

Qy      22 NDAAAEALNAS----------------SKEEALQIIREKIPEKYLFQFHNLNSNLDR 62
           |:| || || :                 :|::| :: : :|||| :  |:  |:| |:
Db      93 NEAKAEKLNQTPEFKAMMEAVKKQALVEFWAKKQAEEVKKIQIPEKEMQDFY--NANKDQ 150

Qy      63 IFDK 66
           :| |
Db     151 LFVK 154
```

Search completed: October 10, 2002, 02:09:53
Job time: 1351 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:48:45 ; Search time 69.63 Seconds
(without alignments)
96.600 Million cell updates/sec

Title: US-09-289-346A-9
Perfect score: 362
Sequence: 1 TLVWGEFQVAGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	356	98.3	352	1	QQCVL1	AL1 protein - toma
2	256	70.7	361	1	QQCVPT	AL1 protein - pota
3	243	67.1	358	2	S07594	hypothetical prote
4	236	65.2	362	1	JQ1887	AL1 protein - toma
5	229	63.3	359	2	S39211	gene C1 protein -
6	225	62.2	349	2	JQ2300	replicase - pepper
7	225	62.2	349	2	S31875	AL1 protein - pepp
8	222	61.3	359	2	S22593	hypothetical prote
9	219	60.5	351	2	JQ2327	AL1 protein - Indi
10	219	60.5	355	1	QQCVW1	AV1 protein - abut
11	216	59.7	358	1	JQ1870	AL1 protein - toma
12	216	59.7	359	2	S39235	gene C1 protein -
13	216	59.7	385	2	S28360	AL1 protein - beet
14	209	57.7	360	2	S59885	replication-associ
15	207	57.2	357	1	QQCVC1	AL1 protein - toma
16	142	39.2	131	2	S45059	AC1 protein (clone
17	125	34.5	347	1	QQCVS1	AL1 protein - squa
18	68	18.8	587	2	JC1419	Fc gamma (IgG) rec
19	67	18.5	840	2	T36175	probable large ATP
20	64.5	17.8	299	2	B71967	probable peptidyl-
21	63.5	17.5	642	2	D90558	topoisomerase iv s
22	63	17.4	714	2	C95382	probable ferrichro
23	62.5	17.3	447	2	T12544	hypothetical prote
24	62	17.1	154	2	AD3475	nitrogen regulator
25	62	17.1	181	2	G97976	conserved hypothet
26	62	17.1	295	2	D42452	C1 protein - tobac
27	62	17.1	340	2	B97011	probably O-sialogl
28	61.5	17.0	201	2	A81380	probable flagellar
29	61.5	17.0	481	2	A70091	probable phosphoes
30	61	16.9	1502	1	RGBYH1	CYC1/CYP3 transcri
31	60.5	16.7	1713	2	A55347	adhesive ligand ep
32	60.5	16.7	2137	2	T05244	hypothetical prote
33	60	16.6	160	2	G82060	hypothetical prote
34	60	16.6	319	2	B90067	L-lactate dehydrog
35	60	16.6	447	2	S52437	CDP-diacylglycerol
36	60	16.6	988	2	F86316	protein T10O22.13
37	59.5	16.4	1033	2	E97700	hypothetical prote
38	59	16.3	338	2	AG3582	iron(III)-binding
39	59	16.3	397	2	B71078	probable NADH oxid
40	59	16.3	416	2	A82892	hypothetical prote
41	59	16.3	862	2	H82182	conserved hypothet
42	59	16.3	1287	2	S55954	viral mRNA transla
43	59	16.3	1341	2	JG0166	LaMDR1 protein - L
44	58.5	16.2	180	2	D84082	hypothetical prote
45	58.5	16.2	230	2	S46332	homeotic protein G

ALIGNMENTS

RESULT 1
QQCVL1
AL1 protein - tomato golden mosaic virus
C;Species: tomato golden mosaic virus
A;Note: host Nicotiana sp. (tobacco)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C;Accession: A04170
R;Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.
EMBO J. 3, 2197-2205, 1984
A;Title: Complete nucleotide sequence of the infectious cloned DNA components of toma
A;Reference number: A04163
A;Accession: A04170
A;Molecule type: DNA
A;Residues: 1-352 <HAM>
C;Comment: The genome consists of two circular, single-stranded DNA components, DNA A
.
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 98.3%; Score 356; DB 1; Length 352;
Best Local Similarity 98.6%; Pred. No. 2.5e-32;
Matches 69; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
QQCVPT
AL1 protein - potato yellow mosaic virus (isolate Venezuela)
C;Species: potato yellow mosaic virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C;Accession: JU0364
R;Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.
J. Gen. Virol. 72, 1515-1520, 1991
A;Title: The nucleotide sequence of the infectious cloned DNA components of potato ye
A;Reference number: JU0362; MUID:91311403
A;Accession: JU0364
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-361 <COU>
A;Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             70.7%;  Score 256;  DB 1;  Length 361;
  Best Local Similarity   69.6%;  Pred. No. 4.5e-21;
  Matches   48;  Conservative   11;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||: ||||||| || |||||||||: :|| |::||:||:|||:|||: ||
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy       61 DRIFDKTPE 69
            |||| | ||
Db     170 DRIFMKAPE 178


RESULT   3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C;Species: cassava latent virus
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C;Accession: S07594
R;Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A;Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A;Reference number: S07590; MUID:90174930
A;Accession: S07594
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <MOR>
A;Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C;Genetics:
A;Map position: segment DNA1
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             67.1%;  Score 243;  DB 2;  Length 358;
  Best Local Similarity   62.9%;  Pred. No. 1.3e-19;
  Matches   44;  Conservative   13;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||: ||||||| |::||| |:||: || ||| :||| :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy       61 DRIFDKTPEP 70
            |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   4
JQ1887
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1887
R;Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A;Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A;Reference number: JQ1885; MUID:93139778
A;Accession: JQ1887
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-362 <DRY>
A;Cross-references: GB:S53251
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             65.2%;  Score 236;  DB 1;  Length 362;
  Best Local Similarity   55.3%;  Pred. No. 8.1e-19;
  Matches   47;  Conservative    9;  Mismatches   13;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || ||||| |||||||| |::||| |:|||  || ||| ::||  |: |: |||||||||

Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy       61 DRI----------------FDKTPE 69
            |||                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   5
S39211
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C;Accession: S39211
R;Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A;Description: High similarity among the tomato yellow leaf curl virus isolates from
A;Reference number: S39209
A;Accession: S39211
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <NOR>
A;Cross-references: EMBL:Z25751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             63.3%;  Score 229;  DB 2;  Length 359;
  Best Local Similarity   60.9%;  Pred. No. 4.9e-18;
  Matches   42;  Conservative    9;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||: ||||||| ||:||| |:|:|| || ||| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy       62 RIFDKTPEP 70
            |:|   | |
Db     171 RVFQVPPAP 179


RESULT   6
JQ2300
replicase - pepper huasteco virus (component A)
N;Alternate names: ORF AL1 protein
C;Species: pepper huasteco virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C;Accession: JQ2300
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A;Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A;Reference number: JQ2299; MUID:94015007
A;Accession: JQ2300
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             62.2%;  Score 225;  DB 2;  Length 349;
  Best Local Similarity   60.0%;  Pred. No. 1.3e-17;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||: ||||||| |::||  |:|||::| |||||||:|: |: :  |||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy       61 DRIFDKTPEP 70
            :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   7
S31875
AL1 protein - pepper rizado amarillo virus
C;Species: pepper rizado amarillo virus
```

```
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Sep-1999
C;Accession: S31875
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera-Bu
submitted to the EMBL Data Library, February 1993
A;Description: Complete nucleotide sequence of pepper huasteco virus: analysis and compa
A;Reference number: S31872
A;Accession: S31875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
A;Note: the source is designated as pepper huasteco virus
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             62.2%;  Score 225;  DB 2;  Length 349;
  Best Local Similarity   60.0%;  Pred. No. 1.3e-17;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||: |||||||  |::||  |:|||::| |||||||:|: |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   8
S22593
hypothetical protein C4 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Sep-1999
C;Accession: S22593
R;Kheyr-Pour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.
Nucleic Acids Res. 19, 6763-6769, 1991
A;Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monoparti
A;Reference number: S22588; MUID:92107660
A;Accession: S22593
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-359 <KHE>
A;Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             61.3%;  Score 222;  DB 2;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 3e-17;
  Matches   40;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||: |||||||  ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   9
JQ2327
AL1 protein - Indian cassava mosaic virus
N;Alternate names: replication-associated protein
C;Species: Indian cassava mosaic virus
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C;Accession: JQ2327; S35883
R;Hong, Y.G.; Robinson, D.J.; Harrison, B.D.
J. Gen. Virol. 74, 2437-2443, 1993
A;Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-tran
A;Reference number: JQ2326; MUID:94065670
A;Accession: JQ2327
A;Molecule type: DNA
A;Residues: 1-351 <HON>
A;Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.5%;  Score 219;  DB 2;  Length 351;
  Best Local Similarity   62.7%;  Pred. No. 6.4e-17;
  Matches   42;  Conservative    9;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       4 WGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           || ||: ||||||| |::||| | |||: || |||:|:||  |  ||  ||:::||||||
Db     113 WGTFQIDGRSARGGQQSANDAYAAALNSGSKSEALKILRELAPRDYLRDFHHISSNLDRI 172

Qy      64 FDKTPEP 70
           | | | |
Db     173 FTKPPPP 179


RESULT  10
QQCVW1
AV1 protein - abutilon mosaic virus (isolate West India)
C;Species: abutilon mosaic virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Apr-1994
C;Accession: A36214
R;Frischmuth, T.; Zimmat, G.; Jeske, H.
Virology 178, 461-468, 1990
A;Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as
A;Reference number: A36214; MUID:91020984
A;Accession: A36214
A;Molecule type: DNA
A;Residues: 1-355 <FRI>
A;Cross-references: EMBL:X15983
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.5%;  Score 219;  DB 1;  Length 355;
  Best Local Similarity   58.6%;  Pred. No. 6.5e-17;
  Matches   41;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |  |||||: ||||||| ||:||: |:||||   : || |::|: |: |: | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
JQ1870
AL1 protein - tomato mottle virus (isolate Florida)
C;Species: tomato mottle virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1870
R;Abouzid, A.M.; Polston, J.E.; Hiebert, E.
J. Gen. Virol. 73, 3225-3229, 1992
A;Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated f
A;Reference number: JQ1869; MUID:93107858
A;Accession: JQ1870
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <ABO>
A;Cross-references: GB:L14460
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.7%;  Score 216;  DB 1;  Length 358;
  Best Local Similarity   55.7%;  Pred. No. 1.4e-16;
  Matches   39;  Conservative   16;  Mismatches   15;  Indels    0;  Gaps    0;
```

```
Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||: ||||||| |::||: |:|||| | : || ::||: |: :: | ||: |||
Db     107 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 166

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     167 ERIFAKAPEP 176
```

RESULT 12
S39235
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Sep-1997
C;Accession: S39235
R;Crespi, S.; Noris, E.; Vaira, A.; Bosco, D.; Accotto, G.
submitted to the EMBL Data Library, December 1993
A;Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.
A;Reference number: S39233
A;Accession: S39235
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <CRE>
A;Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             59.7%;  Score 216;  DB 2;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 1.4e-16;
  Matches   39;  Conservative   12;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||: ||||||| ||:||| |:|:|| || ||| :|:: |  |:  |||::|||| 
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINARSKSEALDVIKQLAPRDYVLHFHNISSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179
```

RESULT 13
S28360
AL1 protein - beet curly top virus
C;Species: beet curly top virus
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
C;Accession: S28360
R;Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
EMBO J. 5, 1761-1767, 1986
A;Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly to
A;Reference number: S28360
A;Accession: S28360
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-385 <STA>
A;Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             59.7%;  Score 216;  DB 2;  Length 385;
  Best Local Similarity   55.7%;  Pred. No. 1.5e-16;
  Matches   39;  Conservative   16;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||: ||||||| ||:||: |:|||||:| ::|||||::|: |: |  | ||| :|
Db     137 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 196

Qy      61 DRIFDKTPEP 70
             :|| : |:|
Db     197 QKIFQRPPDP 206
```

RESULT 14
S59885
replication-associated protein C1 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
C;Accession: S59885
R;Hong, Y.; Harrison, B.D.
submitted to the EMBL Data Library, February 1995
A;Description: Nucleotide sequences from tomato leaf curl viruses from different coun
d geminiviruses.
A;Reference number: S58346
A;Accession: S59885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <HON>
A;Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA88229.1; PID:g974211
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             57.7%;  Score 209;  DB 2;  Length 360;
  Best Local Similarity   60.6%;  Pred. No. 8.7e-16;
  Matches   40;  Conservative   12;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       4 WGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           :| ||: ||||||| |::||| |||:|: ||  || |:||| |: :: ||||||:|||||
Db     113 FGVFQIDGRSARGGQQSANDAYAEAINSGSKAAALDILREKAPKDFVLQFHNLNANLDRI 172

Qy      64 FDKTPE 69
           |  : |
Db     173 FTPSAE 178
```

RESULT 15
QQCVC1
AL1 protein - tomato yellow leaf curl virus
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C;Accession: D40779
R;Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.
Virology 185, 151-161, 1991
A;Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin
A;Reference number: A40779; MUID:92024070
A;Accession: D40779
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-357 <NAV>
A;Cross-references: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             57.2%;  Score 207;  DB 1;  Length 357;
  Best Local Similarity   65.6%;  Pred. No. 1.5e-15;
  Matches   40;  Conservative   10;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       4 WGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           :|  |: ||||||| |::||| |||||: || ||| |::|| |: |: |||||:||||||
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171
```

Search completed: October 10, 2002, 01:48:45
Job time: 563 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:47:19 ; Search time 133.15 Seconds
(without alignments)
58.394 Million cell updates/sec

Title: US-09-289-346A-9
Perfect score: 362
Sequence: 1 TLVWGEFQVAGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	362	100.0	70	21	AAB18685	Mutant peptide der
2	356	98.3	70	21	AAB18677	Peptide fragment f
3	356	98.3	356	21	AAB18687	Amino acid sequenc
4	345	95.3	70	21	AAB18688	Mutant peptide der
5	344	95.0	70	21	AAB18692	Mutant peptide der
6	342	94.5	70	21	AAB18684	Mutant peptide der
7	342	94.5	70	21	AAB18690	Mutant peptide der
8	341	94.2	70	21	AAB18678	Mutant peptide der
9	341	94.2	70	21	AAB18686	Mutant peptide der
10	340	93.9	70	21	AAB18689	Mutant peptide der
11	338	93.4	70	21	AAB18680	Mutant peptide der
12	338	93.4	70	21	AAB18691	Mutant peptide der
13	335	92.5	70	21	AAB18681	Mutant peptide der
14	334	92.3	70	21	AAB18683	Mutant peptide der
15	332	91.7	70	21	AAB18682	Mutant peptide der
16	324	89.5	70	21	AAB18679	Mutant peptide der
17	222	61.3	353	18	AAW34338	Bean golden mosaic
18	222	61.3	353	18	AAW34332	Bean golden mosaic
19	222	61.3	353	18	AAW34333	Bean golden mosaic
20	222	61.3	353	18	AAW34334	Bean golden mosaic
21	222	61.3	353	18	AAW34335	Bean golden mosaic
22	222	61.3	359	17	AAR88870	Sardinian tomato y
23	222	61.3	359	17	AAR88871	Sardinian tomato y
24	222	61.3	359	17	AAR88872	Sardinian tomato y
25	220	60.8	353	8	AAP70407	ORF 4 gene product
26	220	60.8	361	18	AAW34336	Tomato mottle viru
27	220	60.8	361	18	AAW34324	Tomato mottle viru
28	220	60.8	361	18	AAW34325	Tomato mottle viru
29	220	60.8	361	18	AAW34326	Tomato mottle viru
30	213.5	59.0	361	8	AAP70562	Product of ORF 4 f
31	208	57.5	362	19	AAW56495	Tobacco leaf curli
32	207	57.2	357	18	AAW34329	Tomato yellow leaf
33	207	57.2	357	18	AAW34330	Tomato yellow leaf
34	207	57.2	357	18	AAW34331	Tomato yellow leaf
35	199	55.0	357	18	AAW34337	Tomato yellow leaf
36	70.5	19.5	945	22	ABB65231	Drosophila melanog
37	66.5	18.4	665	22	ABB70991	Drosophila melanog
38	65	18.0	292	21	AAY92317	Geminivirus Rep Cl
39	65	18.0	335	21	AAY92318	Mastrevirus Rep in
40	64.5	17.8	299	18	AAW55450	H. pylori ORF 02ae
41	64.5	17.8	512	19	AAW68473	HIV-1 strain YBF30
42	62.5	17.3	447	21	AAY99659	Human GTPase assoc
43	62.5	17.3	447	21	AAY87089	Human secreted pro
44	62.5	17.3	447	21	AAY66744	Membrane-bound pro
45	62.5	17.3	447	21	AAY50947	Human adult aorta

ALIGNMENTS

RESULT 1
AAB18685
ID AAB18685 standard; peptide; 70 AA.
XX
AC AAB18685;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 10
FT /note= "wild type residue replaced with Ala"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DR WPI; 2000-618851/59.

```
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 46; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             100.0%;  Score 362;  DB 21;  Length 70;
  Best Local Similarity   100.0%;  Pred. No. 8.2e-38;
  Matches   70;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 tlvwgefqvagrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   2
AAB18677
ID   AAB18677 standard; peptide; 70 AA.
XX
AC   AAB18677;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Peptide fragment from Rep protein of TGMV (amino acids 110-179).
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Tomato golden mosaic virus.
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 18; 73pp; English.
XX
CC   The present sequence is derived from a geminivirus replication (Rep)
CC   protein, which is also known as AL1. AL1 binds double-stranded DNA,
CC   catalyses cleavage and ligation of single-stranded DNA, and interacts
CC   with other viral and host proteins. Mutants of the AL1 protein are used
CC   to produce transgenic plants. The mutation in AL1 is present in a
CC   ribosome binding region, and expression of mutant AL1 protein imparts
CC   increased resistance to geminivirus infection in the plant. Mutant AL1
CC   proteins are useful for producing plants having increased resistance or
CC   reduced sensitivity to a geminivirus such as tomato golden mosaic virus,
CC   tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl
CC   virus, African cassava mosaic virus, Indian cassava mosaic virus, potato
CC   yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,
CC   squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or
CC   beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             98.3%;  Score 356;  DB 21;  Length 70;
  Best Local Similarity   98.6%;  Pred. No. 4.6e-37;
  Matches   69;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   3
AAB18687
ID   AAB18687 standard; peptide; 356 AA.
XX
AC   AAB18687;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Amino acid sequence of a geminivirus replication protein of TGMV.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 354
FT                   /note= "unspecified amino acid"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
```

```
PS   Disclosure; Page 47-48; 73pp; English.
XX
CC   The present sequence represents a geminivirus replication (Rep)
CC   protein, which is also known as AL1. AL1 binds double-stranded DNA,
CC   catalyses cleavage and ligation of single-stranded DNA, and interacts
CC   with other viral and host proteins. Mutants of the AL1 protein are used
CC   to produce transgenic plants. The mutation in AL1 is present in a
CC   ribosome binding region, and expression of mutant AL1 protein imparts
CC   increased resistance to geminivirus infection in the plant. Mutant AL1
CC   proteins are useful for producing plants having increased resistance or
CC   reduced sensitivity to a geminivirus such as tomato golden mosaic virus,
CC   tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl
CC   virus, African cassava mosaic virus, Indian cassava mosaic virus, potato
CC   yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,
CC   squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or
CC   beet curly top virus.
XX
SQ   Sequence   356 AA;


  Query Match             98.3%;  Score 356;  DB 21;  Length 356;
  Best Local Similarity   98.6%;  Pred. No. 3.6e-36;
  Matches   69;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     110 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 169

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     170 drifdktpep 179


RESULT   4
AAB18688
ID   AAB18688 standard; peptide; 70 AA.
XX
AC   AAB18688;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 19
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 20
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 48; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             95.3%;  Score 345;  DB 21;  Length 70;
  Best Local Similarity   95.7%;  Pred. No. 1.1e-35;
  Matches   67;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||| ||||||||  ||||||||||||||||||||||||||||||||||||||||
Db       1 tlvwgefqvdgrsarggcaasndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT   5
AAB18692
ID   AAB18692 standard; peptide; 70 AA.
XX
AC   AAB18692;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 66
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 69
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
```

```
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 50; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             95.0%;  Score 344;  DB 21;  Length 70;
  Best Local Similarity   95.7%;  Pred. No. 1.5e-35;
  Matches   67;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||| || |
Db      61 drifdatpap 70


RESULT   6
AAB18684
ID   AAB18684 standard; peptide; 70 AA.
XX
AC   AAB18684;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 7
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 8
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             94.5%;  Score 342;  DB 21;  Length 70;
  Best Local Similarity   95.7%;  Pred. No. 2.6e-35;
  Matches   67;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||  | |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 tlvwgeaavdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT   7
AAB18690
ID   AAB18690 standard; peptide; 70 AA.
XX
AC   AAB18690;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 27
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 30
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
```

```
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             94.5%;  Score 342;  DB 21;  Length 70;
  Best Local Similarity   95.7%;  Pred. No. 2.6e-35;
  Matches   67;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||| |||||||||||||| || ||||||||||||||||||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaaalaasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT   8
AAB18678
ID   AAB18678 standard; peptide; 70 AA.
XX
AC   AAB18678;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 12
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 13
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 15
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 42-43; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             94.2%;  Score 341;  DB 21;  Length 70;
  Best Local Similarity   94.3%;  Pred. No. 3.5e-35;
  Matches   66;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||| | :| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1 tlvwgefqvdgaaaaggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT   9
AAB18686
ID   AAB18686 standard; peptide; 70 AA.
XX
AC   AAB18686;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 24
FT                   /note= "wild type residue replaced with Leu"
FT   Misc-difference 25
FT                   /note= "wild type residue replaced with Leu"
FT   Misc-difference 26
FT                   /note= "wild type residue replaced with Leu"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
```

PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DR WPI; 2000-618851/59.
XX
PT Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT -
XX
PS Claim 53; Page 46; 73pp; English.
XX
CC The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
SQ Sequence 70 AA;

Query Match 94.2%; Score 341; DB 21; Length 70;
Best Local Similarity 94.3%; Pred. No. 3.5e-35;
Matches 66; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
||||||||| |||||||||||| ||||||||||||||||||||||||||||||||
Db 1 tlvwgefqvdgrsarggcqtsndlllealnasskeealqiirekipekylfqfhnlnsnl 60

Qy 61 DRIFDKTPEP 70
||||||||||
Db 61 drifdktpep 70

RESULT 10
AAB18689
ID AAB18689 standard; peptide; 70 AA.
XX
AC AAB18689;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 22
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 23
FT /note= "wild type residue replaced with Ala"
XX
PN WO200054573-A1.
XX

PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DR WPI; 2000-618851/59.
XX
PT Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT -
XX
PS Disclosure; Page 48-49; 73pp; English.
XX
CC The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
SQ Sequence 70 AA;

Query Match 93.9%; Score 340; DB 21; Length 70;
Best Local Similarity 95.7%; Pred. No. 4.7e-35;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
||||||||| |||||||||| ||||||||||||||||||||||||||||||||
Db 1 tlvwgefqvdgrsarggcqtsaaaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy 61 DRIFDKTPEP 70
||||||||||
Db 61 drifdktpep 70

RESULT 11
AAB18680
ID AAB18680 standard; peptide; 70 AA.
XX
AC AAB18680;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 42
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 43
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 44

```
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 43-44; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             93.4%;  Score 338;  DB 21;  Length 70;
  Best Local Similarity   94.3%;  Pred. No. 8.3e-35;
  Matches   66;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||| ||||||||||||||||||||||||||||||   ||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiiaaaipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  12
AAB18691
ID   AAB18691 standard; peptide; 70 AA.
XX
AC   AAB18691;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 34
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 35
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 36
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             93.4%;  Score 338;  DB 21;  Length 70;
  Best Local Similarity   94.3%;  Pred. No. 8.3e-35;
  Matches   66;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||| ||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnassaaaalqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  13
AAB18681
ID   AAB18681 standard; peptide; 70 AA.
XX
AC   AAB18681;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
```

```
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 47
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 48
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 49
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 44; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             92.5%;  Score 335;  DB 21;  Length 70;
  Best Local Similarity   94.3%;  Pred. No. 2e-34;
  Matches   66;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||| |||||||||||||||||||||||||||||||||||   ||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipaaalfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  14
AAB18683
ID   AAB18683 standard; peptide; 70 AA.
XX
AC   AAB18683;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 59
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 61
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 62
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             92.3%;  Score 334;  DB 21;  Length 70;
  Best Local Similarity   94.3%;  Pred. No. 2.6e-34;
  Matches   66;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsal 60

Qy      61 DRIFDKTPEP 70
           ||||||||
Db      61 aaifdktpep 70


RESULT  15
AAB18682
ID   AAB18682 standard; peptide; 70 AA.
XX
AC   AAB18682;
XX
```

```
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 52
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 54
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 55
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 44-45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match            91.7%;  Score 332;  DB 21;  Length 70;
  Best Local Similarity  94.3%;  Pred. No. 4.7e-34;
  Matches   66;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||| ||||||||||||||||||||||||||||||||||||| |  |||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfafaalnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


Search completed: October 10, 2002, 01:47:19

Job time: 528 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:02:00 ; Search time 118.38 Seconds
(without alignments)
102.295 Million cell updates/sec

Title: US-09-289-346A-10
Perfect score: 364
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	300	82.4	351	12	Q91R10	Q91r10 tomato seve
2	298	81.9	352	12	Q9E000	Q9e000 tomato rugo
3	281	77.2	226	12	O09727	O09727 leonurus mo
4	281	77.2	226	12	Q9WHF6	Q9whf6 tomato mild
5	281	77.2	361	12	Q67574	Q67574 bean golden
6	279	76.6	225	12	Q9QDB1	Q9qdb1 cowpea gold
7	277	76.1	314	12	Q9ELT8	Q9elt8 sweet potat
8	277	76.1	364	12	Q9QS55	Q9qs55 sweet potat
9	274	75.3	185	12	Q98693	Q98693 sida golden
10	266	73.1	149	12	P88975	P88975 macroptiliu
11	266	73.1	233	12	Q9YLA4	Q9yla4 macroptiliu
12	261	71.7	190	12	Q9Z089	Q9z089 tobacco lea
13	261	71.7	190	12	Q9Z084	Q9z084 tobacco lea
14	260	71.4	190	12	Q9W827	Q9w827 tobacco lea
15	260	71.4	208	12	Q9Z0C4	Q9z0c4 tobacco lea
16	260	71.4	234	12	O39180	O39180 geminivirid
17	260	71.4	359	12	Q91M88	Q91m88 tobacco lea
18	258	70.9	208	12	Q9Z0C0	Q9z0c0 tobacco lea
19	258	70.9	208	12	Q9Z0B8	Q9z0b8 tobacco lea
20	256	70.3	208	12	Q9Z0B6	Q9z0b6 tobacco lea
21	255	70.1	203	12	Q9Z083	Q9z083 tobacco lea
22	251	69.0	363	12	O73577	O73577 cotton leaf
23	250	68.7	363	12	O72705	O72705 cotton leaf
24	250	68.7	363	12	O72719	O72719 cotton leaf
25	249	68.4	349	12	Q88888	Q88888 tomato pseu
26	248	68.1	208	12	Q9Z0A0	Q9z0a0 tobacco lea
27	248	68.1	348	12	Q911W5	Q911w5 macroptiliu
28	247	67.9	190	12	Q9Z086	Q9z086 tobacco lea
29	247	67.9	208	12	Q9Z0C6	Q9z0c6 tobacco lea
30	247	67.9	359	12	Q9YZV4	Q9yzv4 tomato yell
31	247	67.9	359	12	Q9YZV2	Q9yzv2 tomato yell
32	247	67.9	359	12	Q9YUX7	Q9yux7 tomato yell
33	247	67.9	359	12	Q9YL27	Q9yl27 tomato yell
34	246	67.6	359	12	Q88942	Q88942 tomato yell
35	246	67.6	359	12	Q91B86	Q91b86 ageratum ye
36	246	67.6	360	12	Q9DXI0	Q9dxi0 ageratum ye
37	245	67.3	362	12	Q9IN48	Q9in48 okra enatio
38	244	67.0	307	12	Q9IET7	Q9iet7 cotton leaf
39	244	67.0	361	12	O72723	O72723 cotton leaf
40	243	66.8	190	12	Q9Z0A7	Q9z0a7 tobacco lea
41	243	66.8	354	12	Q9IN42	Q9in42 south afric
42	242	66.5	358	12	Q9JEA2	Q9jea2 cassava gem
43	242	66.5	362	12	O56816	O56816 chayote mos
44	241	66.2	231	12	Q96620	Q96620 african tom
45	240	65.9	353	12	O72692	O72692 beet curly

ALIGNMENTS

RESULT 1
Q91R10
ID Q91R10 PRELIMINARY; PRT; 351 AA.
AC Q91R10;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE REP PROTEIN.
GN AC1.
OS Tomato severe rugose virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=158463;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MINAS GERAIS;
RA Rezende W.L., Goulart L.R., Parreira K.S., Figueiredo J.E.F.;
RT "The full-length DNA-A nucleotide sequence of a novel tomato-infecting
RT begomovirus, Tomato severe rugose virus, in Brazil.";
RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY029750; AAK50357.1; -.
SQ SEQUENCE 351 AA; 40122 MW; 87F937A4F873B6CF CRC64;

Query Match 82.4%; Score 300; DB 12; Length 351;
Best Local Similarity 78.6%; Pred. No. 2.5e-26;
Matches 55; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:||||||||||:||   ||||| ||: ||||||||:|||:||||||||||
Db     111 TIEWGEFQIDGRSARGGCQTANDAAAEALNAPSKDVALQIIREKLPEKFLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| : |||
Db     171 DRIFARAPEP 180
```

RESULT 2
Q9E000
ID Q9E000 PRELIMINARY; PRT; 352 AA.

```
AC   Q9E000;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN.
GN   AC1.
OS   Tomato rugose mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=134599;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Fernandes J.J., Fontes E.P.B., Brommonschenkel S.H., Carvalho M.G.,
RA   Zambolim E.M., Zerbini F.M.;
RT   "Molecular Cloning and Characterization of Tomato rugose mosaic virus
RT   (TRMV), a Begomovirus Isolated from Tomatoes at Triangulo Mineiro,
RT   Minas Gerais, Brazil.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF291705; AAG15546.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   352 AA;  40012 MW;  47CD55838E24D613 CRC64;


  Query Match             81.9%;  Score 298;  DB 12;  Length 352;
  Best Local Similarity   78.6%;  Pred. No. 4.2e-26;
  Matches   55;  Conservative    7;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:||||||||||:||   ||||| ||: ||||||||:|||:||||||||||
Db     111 TIEWGEFQIDGRSARGGCQTANDAAAEALNAPSKDIALQIIREKMPEKFLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| : |||
Db     171 DRIFARAPEP 180


RESULT   3
O09727
ID   O09727      PRELIMINARY;      PRT;   226 AA.
AC   O09727;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   REP.
OS   Leonurus mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=58177;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LEMV- BRAZIL 1;
RA   Faria J.C., Maxwell D.P.;
RT   "Variability in geminivirus associated with Phaseolus vulgaris in
RT   Brazil.";
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U92532; AAB51157.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     226    226
SQ   SEQUENCE   226 AA;  25617 MW;  73CDB6E766083FC5 CRC64;


  Query Match             77.2%;  Score 281;  DB 12;  Length 226;
  Best Local Similarity   77.1%;  Pred. No. 2.3e-24;
  Matches   54;  Conservative    4;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||||||||||:||| || ||   ||||| |  |||||:||:|||||||||||||||
Db     111 TVEWGEFQVDGRSSRGGQQTVNDAAAEALNAPDKRTALQIIKEKLPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180


RESULT   4
Q9WHF6
ID   Q9WHF6      PRELIMINARY;      PRT;   226 AA.
AC   Q9WHF6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   tomato mild mottle geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX   NCBI_TaxID=92943;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HN96-H5KW;
RA   Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA   Maxwell D.P.;
RT   "Molecular characterization and DNA-based detection methods for
RT   vegetable-infecting geminiviruses in Central America.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF131071; AAD33471.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     226    226
SQ   SEQUENCE   226 AA;  25941 MW;  2EA4116712871A23 CRC64;


  Query Match             77.2%;  Score 281;  DB 12;  Length 226;
  Best Local Similarity   72.9%;  Pred. No. 2.3e-24;
  Matches   51;  Conservative   12;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| ||:||   ||||||||||::||:||:|:::|||:|||:|||
Db     111 TIEWGEFQIDGRSARGGQQTANDAAAEALNASSKEEAMRIIKEKLPKEFLFQYHNLSSNL 170

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180


RESULT   5
Q67574
ID   Q67574      PRELIMINARY;      PRT;   361 AA.
AC   Q67574;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE  REPLICATIVE PROTEIN.
GN   AL1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,
RA   Maxwell D.P., Russell D.R.;
RT   "Cloning of the complete DNA genomes of four bean-infecting
RT   geminiviruses and determining their infectivity by electric discharge
RT   particle acceleration.";
RL   Phytopathology 81:980-985(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,
RA   Morales F.J., Maxwell D.P.;
RT   "Differentiation of bean-infecting geminiviruses by nucleic acid
```

```
RT   hybridization probes and aspects of bean golden mosaic in Brazil.";
RL   Plant Dis. 75:336-342(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RT   "Genetic diversity in geminiviruses causing bean golden mosaic
RT   disease: The nucleotide sequence of the infectious cloned DNA
RT   components of a Brazilian isolate of bean golden mosaic geminivirus.";
RL   Submitted (MAY-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; M88686; AAA46312.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   361 AA;  41041 MW;  0094C7ACAF06B788 CRC64;


  Query Match             77.2%;  Score 281;  DB 12;  Length 361;
  Best Local Similarity   77.6%;  Pred. No. 3.9e-24;
  Matches   52;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           || |||||||||| ||:||   |||||||||||:|||:||:|||:|||:|||:|||||
Db     113 WGHFQVDGRSARGGQQTANDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
           | | |:|
Db     173 FTKAPDP 179


RESULT   6
Q9QDB1
ID   Q9QDB1      PRELIMINARY;      PRT;   225 AA.
AC   Q9QDB1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   cowpea golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=69263;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CGMV-BR;
RA   Faria J.C.;
RT   "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT   Brazil.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF188708; AAF06318.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     225     225
SQ   SEQUENCE   225 AA;  25766 MW;  1089CB6BD8D15B5D CRC64;


  Query Match             76.6%;  Score 279;  DB 12;  Length 225;
  Best Local Similarity   77.6%;  Pred. No. 3.8e-24;
  Matches   52;  Conservative    7;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           || ||:||||||||| || ||   |||||||||||:|||:|||:|||:|||:||||||
Db     113 WGHFQIDGRSARGGQQTINDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
           | | |||
Db     173 FKKPPEP 179


RESULT   7
Q9ELT8
ID   Q9ELT8      PRELIMINARY;      PRT;   314 AA.
AC   Q9ELT8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION ASSCCIATION PROTEIN.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF288227; AAG01006.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   314 AA;  35153 MW;  6B6220613046943F CRC64;


  Query Match             76.1%;  Score 277;  DB 12;  Length 314;
  Best Local Similarity   79.4%;  Pred. No. 9.5e-24;
  Matches   54;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||||||||||||||| ||:||   ||||| ||| ||||||||:|||||||||| |||
Db     110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy      61 DRIFDKTP 68
           ||||   |
Db     170 DRIFSPPP 177


RESULT   8
Q9QS55
ID   Q9QS55      PRELIMINARY;      PRT;   364 AA.
AC   Q9QS55;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION INITIATION PROTEIN AC1.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A.;
RT   "Cloning of a DNA-A-like genomic component of sweet potato leaf curl
RT   virus : nucleotide sequence and phylogenetic relationships.";
RL   Online Publication.
DR   EMBL; AF104036; AAD47173.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   364 AA;  40680 MW;  5F79752431A09D6E CRC64;
```

```
  Query Match           76.1%;  Score 277;  DB 12;  Length 364;
  Best Local Similarity   79.4%;  Pred. No. 1.1e-23;
  Matches   54;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||||||||||||||| ||:||   ||||| ||| |||||||:||||||||||| |||
Db    110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy     61 DRIFDKTP 68
          ||||   |
Db    170 DRIFSPPP 177


RESULT   9
Q98693
ID   Q98693      PRELIMINARY;      PRT;   185 AA.
AC   Q98693;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   AC1.
OS   sida golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51034;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA;
RA   Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RT   "Genetic Diversity among geminiviruses associated with the weed
RT   species Sida spp, Macroptilium lathyroides, and Wissadula amplissima
RT   from Jamaica.";
RL   Plant Dis. 81:1251-1258(1997).
DR   EMBL; U67926; AAB97865.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     185    185
SQ   SEQUENCE   185 AA;  20975 MW;  3913850A025A5EE1 CRC64;


  Query Match           75.3%;  Score 274;  DB 12;  Length 185;
  Best Local Similarity   70.0%;  Pred. No. 1.1e-23;
  Matches   49;  Conservative   12;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: || ||:|||||||| ||:||   |||||: :||:||:|||||:||:||||:|||:||:
Db     89 TIEWGVFQIDGRSARGGQQTANDAAAEALNSGTKEDALKIIREKLPERYLFQYHNLSSNI 148

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    149 DRIFSKPPEP 158


RESULT  10
P88975
ID   P88975      PRELIMINARY;      PRT;   149 AA.
AC   P88975;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   AC1.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICAN;
RA   Roye M.E.;
RL   Thesis (1996), Biochemistry, University of the West Indies, Jamaica.
DR   EMBL; U75278; AAB36919.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     149    149
SQ   SEQUENCE   149 AA;  16785 MW;  E4CF5EED4C9CD508 CRC64;


  Query Match           73.1%;  Score 266;  DB 12;  Length 149;
  Best Local Similarity   68.6%;  Pred. No. 7.4e-23;
  Matches   48;  Conservative   12;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: || ||:|||||||| |||||   ||||: :|| |::|::||:|||:|||:|||:|||
Db     52 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 111

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    112 DRIFMKDPEP 121


RESULT  11
Q9YLA4
ID   Q9YLA4      PRELIMINARY;      PRT;   233 AA.
AC   Q9YLA4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E.;
RT   "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
RT   from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E., McLaughlin W.A., Maxwell D.P.;
RT   "Molecular characterization of two distinct geminiviruses infecting M.
RT   lathyroides from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF098940; AAD17850.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     233    233
SQ   SEQUENCE   233 AA;  26356 MW;  AA490AF4D2166A02 CRC64;


  Query Match           73.1%;  Score 266;  DB 12;  Length 233;
  Best Local Similarity   68.6%;  Pred. No. 1.2e-22;
  Matches   48;  Conservative   12;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: || ||:|||||||| |||||   ||||: :|| |::|::||:|||:|||:|||:|||
Db    110 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 169

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    170 DRIFMKDPEP 179


RESULT  12
Q9Z089
```

```
ID   Q9Z089      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z089;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA3-1, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001315; BAA34033.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21432 MW;  AAC093D1D1610FAD CRC64;


  Query Match             71.7%;  Score 261;  DB 12;  Length 190;
  Best Local Similarity   60.0%;  Pred. No. 3.7e-22;
  Matches   51;  Conservative    8;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || || ||:||||||||||| :||   |||||||| ||| |||||:|: ::||:|||||||
Db      85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  13
Q9Z084
ID   Q9Z084      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z084;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA5-2, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA5;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001318; BAA34039.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  AAC1C2943E3F01AD CRC64;


  Query Match             71.7%;  Score 261;  DB 12;  Length 190;
  Best Local Similarity   60.0%;  Pred. No. 3.7e-22;
  Matches   51;  Conservative    8;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || || ||:||||||||||| :||   |||||||| ||| |||||:|: ::||:|||||||
Db      85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  14
Q9W827
ID   Q9W827      PRELIMINARY;      PRT;   190 AA.
AC   Q9W827;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE GORA1-1, PARTIAL AND
DE   COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GORA1;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001303; BAA34010.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  93C3742A8EBDB7EB CRC64;


  Query Match             71.4%;  Score 260;  DB 12;  Length 190;
  Best Local Similarity   60.0%;  Pred. No. 4.8e-22;
  Matches   51;  Conservative    8;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || || ||||||||||||| :||   |||||:|| ||| |||||:|: ::||:|||||||
Db      85 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFSNSSFDQVPE 169


RESULT  15
Q9Z0C4
ID   Q9Z0C4      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0C4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE ABURA3-1, PARTIAL AND
DE   COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ABURA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001294; BAA33992.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
```

```
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23526 MW;  249CC31D8729C72D CRC64;


  Query Match             71.4%;  Score 260;  DB 12;  Length 208;
  Best Local Similarity   60.0%;  Pred. No. 5.3e-22;
  Matches   51;  Conservative    8;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || || |||||||||||||| :||   |||||:|| ||| |||||:|: ::|||:|||||||
Db     102 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSNSSFDQVPE 186


Search completed: October 10, 2002, 02:02:00
Job time: 1093 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:09:53 ; Search time 34.62 Seconds
(without alignments)
78.289 Million cell updates/sec

Title: US-09-289-346A-10
Perfect score: 364
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	349	95.9	352	1	VAL1_TGMV	P03567 tomato gold
2	249	68.4	361	1	VAL1_PYMVV	P27258 potato yell
3	242	66.5	358	1	VAL1_CLVK	P14982 cassava lat
4	242	66.5	358	1	VAL1_CLVN	P14972 cassava lat
5	235	64.6	362	1	VAL1_TYLCA	P36279 tomato yell
6	228	62.6	349	1	VAL1_PHUV	Q06923 pepper huas
7	228	62.6	359	1	VAL1_TYLCU	P38609 tomato yell
8	221	60.7	353	1	VAL1_BGMV	P05175 bean golden
9	221	60.7	359	1	VAL1_TYLCM	P27260 tomato yell
10	220	60.4	355	1	VAL1_ABMVW	P21947 abutilon mo
11	217	59.6	358	1	VAL1_BCTV	P14991 beet curly
12	217	59.6	361	1	VAL1_TMOV	Q06657 tomato mott
13	206	56.6	357	1	VAL1_TYLCV	P27259 tomato yell
14	125	34.3	347	1	VAL1_SLCV	P29048 squash leaf
15	63	17.3	342	1	RSMC_ECOLI	P39406 escherichia
16	62	17.0	295	1	VAL1_TYDVA	P31617 tobacco yel
17	62	17.0	387	1	Y4PF_RHISN	P55615 rhizobium s
18	61	16.8	555	1	GLPD_BACSU	P18158 bacillus su
19	61	16.8	836	1	GCSR_HUMAN	Q99062 homo sapien
20	60.5	16.6	447	1	TBL2_HUMAN	Q9y4p3 homo sapien
21	60	16.5	470	1	RHBA_RHIME	Q9z3r2 rhizobium m
22	60	16.5	509	1	C4A3_DROME	Q9vms7 drosophila
23	60	16.5	799	1	AFSK_STRCO	P54741 streptomyce
24	60	16.5	807	1	AFSK_STRGR	P54742 streptomyce
25	60	16.5	1117	1	TERT_TETTH	O77448 tetrahymena
26	59.5	16.3	79	1	AKHD_DROME	P17975 drosophila
27	58.5	16.1	323	1	VAL1_PASVK	Q00338 panicum str
28	58.5	16.1	397	1	METL_RAT	P13444 rattus norv
29	58.5	16.1	1610	1	CCAD_MESAU	Q99244 mesocricetu
30	58.5	16.1	2161	1	CCAD_HUMAN	Q01668 homo sapien
31	58.5	16.1	2203	1	CCAD_RAT	P27732 rattus norv
32	58	15.9	256	1	YKJ9_YEAST	P34247 saccharomyc
33	58	15.9	354	1	CARA_METJA	Q58425 methanococc
34	58	15.9	543	1	NFL_HUMAN	P07196 homo sapien
35	58	15.9	1287	1	SKI2_YEAST	P35207 saccharomyc
36	57	15.7	784	1	LON_RICPR	Q9zd92 rickettsia
37	57	15.7	970	1	DA81_YEAST	P21657 saccharomyc
38	57	15.7	1517	1	RPOC_CAMJE	Q9pi30 campylobact
39	56.5	15.5	266	1	ETFB_MYCLE	O33095 mycobacteri
40	56.5	15.5	266	1	ETFB_MYCTU	O53276 mycobacteri
41	56.5	15.5	395	1	METL_HUMAN	Q00266 homo sapien
42	56.5	15.5	1026	1	MY1B_DROME	Q23979 drosophila
43	56.5	15.5	1042	1	T1RH_METJA	Q60295 methanococc
44	56	15.4	247	1	YCP4_YEAST	P25349 saccharomyc
45	56	15.4	299	1	Y175_HELPY	P56112 helicobacte

ALIGNMENTS

RESULT 1
VAL1_TGMV
ID VAL1_TGMV STANDARD; PRT; 352 AA.
AC P03567;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein.
GN AC1.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10831;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;
RT "Complete nucleotide sequence of the infectious cloned DNA components
RT of tomato golden mosaic virus: potential coding regions and regulatory
RT sequences.";
RL EMBO J. 3:2197-2205(1984).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; K02029; -; NOT_ANNOTATED_CDS.
DR PIR; A04170; QQCVL1.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR ProDom; PD000736; Gemini_AL1; 1.
KW ATP-binding.
FT NP_BIND 223 230 ATP (POTENTIAL).
SQ SEQUENCE 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;

Query Match 95.9%; Score 349; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 1.7e-31;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
VAL1_PYMVV
ID VAL1_PYMVV STANDARD; PRT; 361 AA.

```
AC   P27258;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 protein.
OS   Potato yellow mosaic virus (isolate Venezuela).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10828;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91311403; PubMed=1856690;
RA   Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT   "The nucleotide sequence of the infectious cloned DNA components of
RT   potato yellow mosaic virus.";
RL   J. Gen. Virol. 72:1515-1520(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; D00940; BAA00782.1; -.
DR   PIR; JU0364; QQCVPT.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   361 AA;  40850 MW;  5627A33BF1264383 CRC64;


  Query Match             68.4%;  Score 249;  DB 1;  Length 361;
  Best Local Similarity   66.7%;  Pred. No. 2.3e-20;
  Matches   46;  Conservative   11;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:||||||||| || ||   ||||: :|| |::|||:||:|||:|||:|||: ||
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 DRIFDKTPE 69
           |||| | ||
Db     170 DRIFMKAPE 178


RESULT   3
VAL1_CLVK
ID   VAL1_CLVK      STANDARD;      PRT;   358 AA.
AC   P14982;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (40.4 kDa protein).
GN   AC1.
OS   Cassava latent virus (strain West Kenyan 844).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10818;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Gay M.R.;
RT   "Nucleotide sequence of cassava latent virus DNA.";
RL   Nature 301:260-262(1983).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; J02057; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40346 MW;  ED173E753EE92D69 CRC64;


  Query Match             66.5%;  Score 242;  DB 1;  Length 358;
  Best Local Similarity   61.4%;  Pred. No. 1.3e-19;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||   :|||: || ||| :||| :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   4
VAL1_CLVN
ID   VAL1_CLVN      STANDARD;      PRT;   358 AA.
AC   P14972;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (40.4 kDa protein).
GN   AC1.
OS   Cassava latent virus (strain Nigerian).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10819;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90174930; PubMed=2308831;
RA   Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT   "Nucleotide sequence of the infectious cloned DNA components of
RT   African cassava mosaic virus (Nigerian strain).";
RL   Nucleic Acids Res. 18:197-198(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X17095; CAA34953.1; -.
DR   PIR; S07594; S07594.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40435 MW;  1DB16BB0CB2D5E2C CRC64;


  Query Match             66.5%;  Score 242;  DB 1;  Length 358;
  Best Local Similarity   61.4%;  Pred. No. 1.3e-19;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||   :|||: || ||| :||| :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168
```

```
Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   5
VAL1_TYLCA
ID   VAL1_TYLCA     STANDARD;      PRT;   362 AA.
AC   P36279;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=36447;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93139778; PubMed=8423446;
RA   Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;
RT   "Nucleotide sequence and genome organization of tomato leaf curl
RT   geminivirus.";
RL   J. Gen. Virol. 74:147-151(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
DR   PIR; JQ1887; JQ1887.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   362 AA;  41197 MW;  343E7184B4704098 CRC64;


  Query Match             64.6%;  Score 235;  DB 1;  Length 362;
  Best Local Similarity   54.1%;  Pred. No. 8.1e-19;
  Matches   46;  Conservative    9;  Mismatches   14;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || |||||:|||||||| |::||   :|||  || ||| ::||  |: |: |||||||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   6
VAL1_PHUV
ID   VAL1_PHUV      STANDARD;      PRT;   349 AA.
AC   Q06923;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AL1.
OS   Pepper huasteco virus (PHV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=28349;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94015007; PubMed=8409944;
RA   Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA   Rivera-Bustamante R.F.;
RT   "Complete nucleotide sequence of pepper huasteco virus: analysis and
RT   comparison with bipartite geminiviruses.";
RL   J. Gen. Virol. 74:2225-2231(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70418; CAA49856.1; -.
DR   PIR; S31875; S31875.
DR   PIR; JQ2300; JQ2300.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (BY SIMILARITY).
SQ   SEQUENCE   349 AA;  39722 MW;  D5F4E76CD56370F4 CRC64;


  Query Match             62.6%;  Score 228;  DB 1;  Length 349;
  Best Local Similarity   60.0%;  Pred. No. 4.7e-18;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| |::||   :|||::| |||||||:|: |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   7
VAL1_TYLCU
ID   VAL1_TYLCU     STANDARD;      PRT;   359 AA.
AC   P38609;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=37139;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94256836; PubMed=8198442;
RA   Noris E., Hidalgo E., Accotto G., Moriones E.;
RT   "High similarity among the tomato yellow leaf curl virus isolates
RT   from the west Mediterranean basin: the nucleotide sequence of an
RT   infectious clone from Spain.";
RL   Arch. Virol. 135:165-170(1994).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z25751; CAA81026.1; -.
DR   PIR; S39211; S39211.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  41065 MW;  2D170A51EF80A3EC CRC64;


  Query Match             62.6%;  Score 228;  DB 1;  Length 359;
```

```
  Best Local Similarity   59.4%;  Pred. No. 4.8e-18;
  Matches   41;  Conservative    9;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||   :|:|| || ||| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
             |:|   | |
Db     171 RVFQVPPAP 179


RESULT   8
VAL1_BGMV
ID   VAL1_BGMV      STANDARD;      PRT;   353 AA.
AC   P05175;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   AL1 protein (40.2 kDa protein).
GN   AC1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Howarth A.J., Caton J., Bossert M., Goodman R.M.;
RT   "Nucleotide sequence of bean golden mosaic virus and a model for gene
RT   regulation in geminiviruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M10070; AAA46318.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   353 AA;  40190 MW;  80FA779DF6029A34 CRC64;


  Query Match             60.7%;  Score 221;  DB 1;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 2.8e-17;
  Matches   42;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:|||||||||| |::||   :|||| | | || |::|: |: |: | ||: |||
Db     110 TIEWGQFQVDGRSARGGQQSANDSYAKALNADSIESALTILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFFKVPEP 179


RESULT   9
VAL1_TYLCM
ID   VAL1_TYLCM     STANDARD;      PRT;   359 AA.
AC   P27260;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10833;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92107660; PubMed=1840676;
RA   Kheyr-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,
RA   Gronenborn B.;
RT   "Tomato yellow leaf curl virus from Sardinia is a
RT   whitefly-transmitted monopartite geminivirus.";
RL   Nucleic Acids Res. 19:6763-6769(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X61153; CAA43466.1; -.
DR   PIR; S22593; S22593.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  40733 MW;  9717B4A07C93EFA7 CRC64;


  Query Match             60.7%;  Score 221;  DB 1;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 2.9e-17;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||   :|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|  |.|
Db     171 KVFQVPPAP 179


RESULT  10
VAL1_ABMVW
ID   VAL1_ABMVW     STANDARD;      PRT;   355 AA.
AC   P21947;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AC1.
OS   Abutilon mosaic virus (isolate West India).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10816;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91020984; PubMed=2219703;
RA   Frischmuth T., Zimmat G., Jeske H.;
RT   "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT   as well as eukaryotic features.";
RL   Virology 178:461-468(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR   PIR; A36214; QQCVW1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   355 AA;  40257 MW;  16A2CA8A63251E95 CRC64;


  Query Match             60.4%;  Score 220;  DB 1;  Length 355;
  Best Local Similarity   58.6%;  Pred. No. 3.7e-17;
  Matches   41;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |  |||||:|||||||| ||:||   :||||   : || |::|: |: |: | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
VAL1_BCTV
ID   VAL1_BCTV      STANDARD;      PRT;   358 AA.
AC   P14991;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 protein (40.8 kDa protein).
OS   Beet curly top virus (BCTV).
OC   Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
OX   NCBI_TaxID=10840;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT   "The nucleotide sequence of an infectious clone of the geminivirus
RT   beet curly top virus.";
RL   EMBO J. 5:1761-1767(1986).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40889 MW;  39A45FE3C0B9C333 CRC64;


  Query Match             59.6%;  Score 217;  DB 1;  Length 358;
  Best Local Similarity   55.7%;  Pred. No. 8e-17;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| ||:||   :||||:| ::||||::|: |: |   | ||| :|
Db     110 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 169

Qy      61 DRIFDKTPEP 70
           :|| : |:|
Db     170 QKIFQRPPDP 179


RESULT  12
VAL1_TMOV
ID   VAL1_TMOV      STANDARD;      PRT;   361 AA.
AC   Q06657;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AL1.
OS   Tomato mottle virus (isolate Florida) (TMoV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=36449;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93107858; PubMed=1469361;
RA   Abouzid A.M., Polston J.E., Hiebert E.;
RT   "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT   isolated from tomatoes in Florida.";
RL   J. Gen. Virol. 73:3225-3229(1992).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L14460; AAC32414.1; -.
DR   PIR; JQ1870; JQ1870.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (BY SIMILARITY).
SQ   SEQUENCE   361 AA;  40516 MW;  8138B65CEEAC6950 CRC64;


  Query Match             59.6%;  Score 217;  DB 1;  Length 361;
  Best Local Similarity   55.7%;  Pred. No. 8.1e-17;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||   :|||| | : || ::||: |: :: | ||: |||
Db     110 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  13
VAL1_TYLCV
ID   VAL1_TYLCV     STANDARD;      PRT;   357 AA.
AC   P27259;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10832;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92024070; PubMed=1926771;
RA   Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;
RT   "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RT   with a single genomic component.";
RL   Virology 185:151-161(1991).
```

```
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15656; CAA33688.1; -.
DR   PIR; D40779; QQCVC1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     219    226       ATP (POTENTIAL).
SQ   SEQUENCE   357 AA;  40678 MW;  939AB68E1AB3B2A7 CRC64;


  Query Match             56.6%;  Score 206;  DB 1;  Length 357;
  Best Local Similarity   63.9%;  Pred. No. 1.3e-15;
  Matches   39;  Conservative   10;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           :|  |:|||||||| |::||   ||||: || ||| |::|| |: |: |||||:||||||
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171


RESULT  14
VAL1_SLCV
ID   VAL1_SLCV      STANDARD;      PRT;   347 AA.
AC   P29048;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   AL1 protein.
OS   Squash leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10829;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91082449; PubMed=1984668;
RA   Lazarowitz S.G., Lazdins I.B.;
RT   "Infectivity and complete nucleotide sequence of the cloned genomic
RT   components of a bipartite squash leaf curl geminivirus with a broad
RT   host range phenotype.";
RL   Virology 180:58-69(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M38183; AAC32410.1; ALT_INIT.
DR   PIR; C36785; QQCVS1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     218    225       ATP (POTENTIAL).
SQ   SEQUENCE   347 AA;  39110 MW;  AFDAEBDDE122110E CRC64;


  Query Match             34.3%;  Score 125;  DB 1;  Length 347;
  Best Local Similarity   37.9%;  Pred. No. 1.3e-06;
  Matches   25;  Conservative   15;  Mismatches   22;  Indels    4;  Gaps    1;

Qy       5 GEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRIF 64
           |:::| |    |     :|:   |:|| |  ||| ||:   |: ::  :||| :|::|:|
Db     116 GQYKVSG----GSKSNKDDVYHNAVNAGSAGEALDIIKAGDPKTFIVNYHNLLANVERLF 171

Qy      65 DKTPEP 70
           | |||
Db     172 QKPPEP 177


RESULT  15
RSMC_ECOLI
ID   RSMC_ECOLI     STANDARD;      PRT;   342 AA.
AC   P39406;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ribosomal RNA small subunit methyltransferase C (EC 2.1.1.52) (rRNA
DE   (guanine-N2-)-methyltransferase) (16S rRNA m2G1207 methyltransferase).
GN   RSMC OR B4371.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=95334362; PubMed=7610040;
RA   Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA   Blattner F.R.;
RT   "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT   region from 92.8 through 100 minutes.";
RL   Nucleic Acids Res. 23:2105-2119(1995).
RN   [2]
RP   SEQUENCE OF 1-19, AND CHARACTERIZATION.
RX   MEDLINE=99091632; PubMed=9873033;
RA   Tscherne J.S., Nurse K., Popienick P., Ofengand J.;
RT   "Purification, cloning, and characterization of the 16S RNA m2G1207
RT   methyltransferase from Escherichia coli.";
RL   J. Biol. Chem. 274:924-929(1999).
CC   -!- FUNCTION: SPECIFICALLY METHYLATES THE GUANOSINE IN POSITION 1207
CC       OF 16S RRNA IN THE 30S PARTICLE.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + rRNA = S-adenosyl-L-
CC       homocysteine + rRNA containing N2-methylguanine.
CC   -!- COFACTOR: REQUIRES MAGNESIUM FOR ACTIVITY.
CC   -!- SIMILARITY: BELONGS TO THE METHYLTRANSFERASE SUPERFAMILY. RSMC
CC       SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U14003; AAA97267.1; -.
DR   EMBL; AE000507; AAC77324.1; -.
DR   EcoGene; EG12596; rsmC.
DR   InterPro; IPR002052; N6_Mtase.
DR   InterPro; IPR000051; SAM_bind.
DR   PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
KW   rRNA processing; Transferase; Methyltransferase; Magnesium;
KW   Complete proteome.
FT   INIT_MET      0      0
SQ   SEQUENCE   342 AA;  37493 MW;  C7A318155700302D CRC64;


  Query Match             17.3%;  Score 63;  DB 1;  Length 342;
```

Best Local Similarity 31.1%; Pred. No. 9.5;
Matches 14; Conservative 7; Mismatches 16; Indels 8; Gaps 1;

```
Qy       4 WGEFQVDGRS--------ARGGCQTSNDLLLEALNASSKEEALQI 40
           |||: ||| :         :| |     : |||  |    :| : | :
Db     158 WGEYSVDGLTVKTLPGVFSRDGLDVGSQLLLSTLTPHTKGKVLDV 202
```

Search completed: October 10, 2002, 02:09:53
Job time: 1351 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:48:45 ; Search time 69.63 Seconds
(without alignments)
96.600 Million cell updates/sec

Title: US-09-289-346A-10
Perfect score: 364
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	349	95.9	352	1	QQCVL1	AL1 protein - toma
2	249	68.4	361	1	QQCVPT	AL1 protein - pota
3	242	66.5	358	2	S07594	hypothetical prote
4	235	64.6	362	1	JQ1887	AL1 protein - toma
5	228	62.6	349	2	JQ2300	replicase - pepper
6	228	62.6	349	2	S31875	AL1 protein - pepp
7	228	62.6	359	2	S39211	gene C1 protein -
8	221	60.7	359	2	S22593	hypothetical prote
9	220	60.4	355	1	QQCVW1	AV1 protein - abut
10	218	59.9	351	2	JQ2327	AL1 protein - Indi
11	217	59.6	358	1	JQ1870	AL1 protein - toma
12	217	59.6	385	2	S28360	AL1 protein - beet
13	215	59.1	359	2	S39235	gene C1 protein -
14	208	57.1	360	2	S59885	replication-associ
15	206	56.6	357	1	QQCVC1	AL1 protein - toma
16	138	37.9	131	2	S45059	AC1 protein (clone
17	125	34.3	347	1	QQCVS1	AL1 protein - squa
18	67	18.4	714	2	C95382	probable ferrichro
19	64	17.6	154	2	AD3475	nitrogen regulator
20	64	17.6	587	2	JC1419	Fc gamma (IgG) rec
21	63.5	17.4	1229	2	T48959	kinesin-like prote
22	63	17.3	343	2	A98295	ribosomal RNA smal
23	63	17.3	343	2	C86136	probable enzyme yj
24	63	17.3	343	2	S56595	rRNA (guanine-N2-)
25	62	17.0	295	2	D42452	C1 protein - tobac
26	62	17.0	329	2	D85956	hypothetical prote
27	62	17.0	329	2	A91111	hypothetical prote
28	62	17.0	416	2	A82892	hypothetical prote
29	61.5	16.9	333	2	B42476	hypothetical prote
30	61.5	16.9	333	2	AC0766	probable glycosylt
31	61.5	16.9	481	2	A70091	probable phosphoes
32	61.5	16.9	1279	2	E64709	type IIS restricti
33	61	16.8	555	2	C45868	glycerol-3-phospha
34	61	16.8	771	2	B38252	granulocyte colony
35	61	16.8	783	2	JH0329	granulocyte colony
36	61	16.8	863	2	C38252	granulocyte colony
37	60.5	16.6	447	2	T12544	hypothetical prote
38	60.5	16.6	1792	2	T13939	myosin V - fruit f
39	60	16.5	247	2	H86844	glutamine ABC tran
40	60	16.5	470	2	T46814	gamma-aminobutyrat
41	60	16.5	470	2	B95419	diaminobutyrate--p
42	60	16.5	799	2	T48889	serine/threonine p
43	60	16.5	1117	2	T14891	telomerase (EC 2.7
44	60	16.5	1265	2	T02131	hypothetical prote
45	59.5	16.3	79	1	A58656	adipokinetic hormo

ALIGNMENTS

RESULT 1
QQCVL1
AL1 protein - tomato golden mosaic virus
C;Species: tomato golden mosaic virus
A;Note: host Nicotiana sp. (tobacco)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C;Accession: A04170
R;Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.
EMBO J. 3, 2197-2205, 1984
A;Title: Complete nucleotide sequence of the infectious cloned DNA components of toma
A;Reference number: A04163
A;Accession: A04170
A;Molecule type: DNA
A;Residues: 1-352 <HAM>
C;Comment: The genome consists of two circular, single-stranded DNA components, DNA A
.
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 95.9%; Score 349; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 2.6e-31;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db    111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db    171 DRIFDKTPEP 180
```

RESULT 2
QQCVPT
AL1 protein - potato yellow mosaic virus (isolate Venezuela)
C;Species: potato yellow mosaic virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C;Accession: JU0364
R;Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.
J. Gen. Virol. 72, 1515-1520, 1991
A;Title: The nucleotide sequence of the infectious cloned DNA components of potato ye
A;Reference number: JU0362; MUID:91311403
A;Accession: JU0364
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-361 <COU>
A;Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             68.4%;  Score 249;  DB 1;  Length 361;
  Best Local Similarity   66.7%;  Pred. No. 4e-20;
  Matches   46;  Conservative   11;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:|||||||| || ||   |||| :|| |::||:||:|||:|||:|||: ||
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 DRIFDKTPE 69
           |||| | ||
Db     170 DRIFMKAPE 178


RESULT   3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C;Species: cassava latent virus
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C;Accession: S07594
R;Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A;Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A;Reference number: S07590; MUID:90174930
A;Accession: S07594
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <MOR>
A;Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C;Genetics:
A;Map position: segment DNA1
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             66.5%;  Score 242;  DB 2;  Length 358;
  Best Local Similarity   61.4%;  Pred. No. 2.4e-19;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||   :|||: || ||| :||| :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   4
JQ1887
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1887
R;Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A;Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A;Reference number: JQ1885; MUID:93139778
A;Accession: JQ1887
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-362 <DRY>
A;Cross-references: GB:S53251
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             64.6%;  Score 235;  DB 1;  Length 362;
  Best Local Similarity   54.1%;  Pred. No. 1.5e-18;
  Matches   46;  Conservative    9;  Mismatches   14;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || |||||:|||||||| |::||   :|||  || ||| ::||  |: |: |||||||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   5
JQ2300
replicase - pepper huasteco virus (component A)
N;Alternate names: ORF AL1 protein
C;Species: pepper huasteco virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C;Accession: JQ2300
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A;Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A;Reference number: JQ2299; MUID:94015007
A;Accession: JQ2300
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             62.6%;  Score 228;  DB 2;  Length 349;
  Best Local Similarity   60.0%;  Pred. No. 8.7e-18;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| |::||   :|||::| |||||||:|: |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   6
S31875
AL1 protein - pepper rizado amarillo virus
C;Species: pepper rizado amarillo virus
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Sep-1999
C;Accession: S31875
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera
submitted to the EMBL Data Library, February 1993
A;Description: Complete nucleotide sequence of pepper huasteco virus: analysis and co
A;Reference number: S31872
A;Accession: S31875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
A;Note: the source is designated as pepper huasteco virus
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             62.6%;  Score 228;  DB 2;  Length 349;
  Best Local Similarity   60.0%;  Pred. No. 8.7e-18;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| |::||   :|||::| |||||||:|: |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   7
S39211
gene C1 protein - tomato yellow leaf curl virus
```

```
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C;Accession: S39211
R;Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A;Description: High similarity among the tomato yellow leaf curl virus isolates from the
A;Reference number: S39209
A;Accession: S39211
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <NOR>
A;Cross-references: EMBL:Z25751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             62.6%;  Score 228;  DB 2;  Length 359;
  Best Local Similarity   59.4%;  Pred. No. 9e-18;
  Matches   41;  Conservative    9;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:||||||||| ||:||   :|:|| || ||| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   8
S22593
hypothetical protein C4 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Sep-1999
C;Accession: S22593
R;Kheyr-Pour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.
Nucleic Acids Res. 19, 6763-6769, 1991
A;Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monoparti
A;Reference number: S22588; MUID:92107660
A;Accession: S22593
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-359 <KHE>
A;Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.7%;  Score 221;  DB 2;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 5.4e-17;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:||||||||| ||:||   :|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   9
QQCVW1
AV1 protein - abutilon mosaic virus (isolate West India)
C;Species: abutilon mosaic virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Apr-1994
C;Accession: A36214
R;Frischmuth, T.; Zimmat, G.; Jeske, H.
Virology 178, 461-468, 1990
A;Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as wel
A;Reference number: A36214; MUID:91020984
A;Accession: A36214
A;Molecule type: DNA
A;Residues: 1-355 <FRI>
A;Cross-references: EMBL:X15983
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.4%;  Score 220;  DB 1;  Length 355;
  Best Local Similarity   58.6%;  Pred. No. 6.9e-17;
  Matches   41;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |  |||||:|||||||| ||:||   :||||   : || |::|: |: |: | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  10
JQ2327
AL1 protein - Indian cassava mosaic virus
N;Alternate names: replication-associated protein
C;Species: Indian cassava mosaic virus
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C;Accession: JQ2327; S35883
R;Hong, Y.G.; Robinson, D.J.; Harrison, B.D.
J. Gen. Virol. 74, 2437-2443, 1993
A;Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-t
A;Reference number: JQ2326; MUID:94065670
A;Accession: JQ2327
A;Molecule type: DNA
A;Residues: 1-351 <HON>
A;Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.9%;  Score 218;  DB 2;  Length 351;
  Best Local Similarity   61.2%;  Pred. No. 1.1e-16;
  Matches   41;  Conservative    9;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           || ||:||||||||| |::||    |||: || |||:|:||  |  ||  ||:::||||||
Db     113 WGTFQIDGRSARGGQQSANDAYAAALNSGSKSEALKILRELAPRDYLRDFHHISSNLDRI 172

Qy      64 FDKTPEP 70
           | | | |
Db     173 FTKPPPP 179


RESULT  11
JQ1870
AL1 protein - tomato mottle virus (isolate Florida)
C;Species: tomato mottle virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1870
R;Abouzid, A.M.; Polston, J.E.; Hiebert, E.
J. Gen. Virol. 73, 3225-3229, 1992
A;Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated f
A;Reference number: JQ1869; MUID:93107858
A;Accession: JQ1870
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <ABO>
A;Cross-references: GB:L14460
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.6%;  Score 217;  DB 1;  Length 358;
  Best Local Similarity   55.7%;  Pred. No. 1.5e-16;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;
```

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||   :|||| | : || ::||: |: :: | ||: |||
Db     107 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 166

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     167 ERIFAKAPEP 176
```

RESULT 12
S28360
AL1 protein - beet curly top virus
C;Species: beet curly top virus
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
C;Accession: S28360
R;Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
EMBO J. 5, 1761-1767, 1986
A;Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly to
A;Reference number: S28360
A;Accession: S28360
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-385 <STA>
A;Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             59.6%;  Score 217;  DB 2;  Length 385;
  Best Local Similarity   55.7%;  Pred. No. 1.7e-16;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| ||:||   :||||:| ::|||||::|: |: | | ||| :|
Db     137 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 196

Qy      61 DRIFDKTPEP 70
           :|| : |:|
Db     197 QKIFQRPPDP 206
```

RESULT 13
S39235
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Sep-1997
C;Accession: S39235
R;Crespi, S.; Noris, E.; Vaira, A.; Bosco, D.; Accotto, G.
submitted to the EMBL Data Library, December 1993
A;Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.
A;Reference number: S39233
A;Accession: S39235
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <CRE>
A;Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             59.1%;  Score 215;  DB 2;  Length 359;
  Best Local Similarity   55.1%;  Pred. No. 2.5e-16;
  Matches   38;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||   :|:|| || ||| :|:: |  |:  |||::||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINARSKSEALDVIKQLAPRDYVLHFHNISSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179
```

RESULT 14
S59885
replication-associated protein C1 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
C;Accession: S59885
R;Hong, Y.; Harrison, B.D.
submitted to the EMBL Data Library, February 1995
A;Description: Nucleotide sequences from tomato leaf curl viruses from different coun
d geminiviruses.
A;Reference number: S58346
A;Accession: S59885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <HON>
A;Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA88229.1; PID:g974211
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             57.1%;  Score 208;  DB 2;  Length 360;
  Best Local Similarity   59.1%;  Pred. No. 1.6e-15;
  Matches   39;  Conservative   12;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           :| ||:|||||||| |::||   ||:|: ||  || |:||| |: :: ||||||:|||||
Db     113 FGVFQIDGRSARGGQQSANDAYAEAINSGSKAAALDILREKAPKDFVLQFHNLNANLDRI 172

Qy      64 FDKTPE 69
           |  : |
Db     173 FTPSAE 178
```

RESULT 15
QQCVC1
AL1 protein - tomato yellow leaf curl virus
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C;Accession: D40779
R;Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.
Virology 185, 151-161, 1991
A;Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin
A;Reference number: A40779; MUID:92024070
A;Accession: D40779
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-357 <NAV>
A;Cross-references: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             56.6%;  Score 206;  DB 1;  Length 357;
  Best Local Similarity   63.9%;  Pred. No. 2.6e-15;
  Matches   39;  Conservative   10;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           :|  |:|||||||| |::||   |||| || ||| |::|| |: |: ||||||:|||||
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171
```

Search completed: October 10, 2002, 01:48:46
Job time: 564 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:09:03 ; Search time 407.82 Seconds
(without alignments)
47.487 Million cell updates/sec

Title: US-09-289-346A-10
Perfect score: 364
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1006125 seqs, 276659714 residues

Total number of hits satisfying chosen parameters: 1006125

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	67	18.4	340	6	US-10-219-999-59588	Sequence 59588, A
2	65.5	18.0	945	5	US-09-791-537-132140	Sequence 132140,
3	65	17.9	293	6	US-10-151-336-3	Sequence 3, Appli
4	65	17.9	335	6	US-10-151-336-6	Sequence 6, Appli
5	64	17.6	587	5	US-09-791-537-33198	Sequence 33198, A
6	64	17.6	712	5	US-09-893-519A-35	Sequence 35, Appl
7	64	17.6	793	6	US-10-179-131-6704	Sequence 6704, Ap
8	63.5	17.4	1229	5	US-09-791-537-142650	Sequence 142650,
9	63	17.3	343	5	US-09-791-537-100065	Sequence 100065,
10	62.5	17.2	1286	5	US-09-791-537-23759	Sequence 23759, A
11	62	17.0	21	1	PCT-US02-21494-17	Sequence 17, Appl
12	62	17.0	21	6	US-10-105-232-17	Sequence 17, Appl
13	62	17.0	329	6	US-10-053-853A-1143	Sequence 1143, Ap
14	62	17.0	631	5	US-09-791-537-2671	Sequence 2671, Ap
15	62	17.0	643	5	US-09-791-537-113157	Sequence 113157,
16	62	17.0	646	5	US-09-791-537-113627	Sequence 113627,
17	62	17.0	658	5	US-09-791-537-2673	Sequence 2673, Ap
18	61.5	16.9	1279	5	US-09-882-227-388	Sequence 388, App
19	61	16.8	177	6	US-10-212-778-1167	Sequence 1167, Ap
20	61	16.8	255	6	US-10-212-778-1478	Sequence 1478, Ap
21	61	16.8	555	5	US-09-791-537-28653	Sequence 28653, A
22	61	16.8	555	7	US-60-360-039-17661	Sequence 17661, A
23	61	16.8	771	5	US-09-791-537-78442	Sequence 78442, A
24	61	16.8	783	5	US-09-791-537-78448	Sequence 78448, A
25	61	16.8	836	5	US-09-791-537-96930	Sequence 96930, A
26	61	16.8	863	5	US-09-791-537-5078	Sequence 5078, Ap
27	60.5	16.6	447	5	US-09-991-150-369	Sequence 369, App
28	60.5	16.6	447	5	US-09-989-328-369	Sequence 369, App
29	60.5	16.6	447	6	US-10-103-295-129	Sequence 129, App
30	60.5	16.6	447	6	US-10-121-062-254	Sequence 254, App
31	60.5	16.6	447	6	US-10-173-689-254	Sequence 254, App
32	60.5	16.6	447	6	US-10-173-690-254	Sequence 254, App
33	60.5	16.6	447	6	US-10-173-691-254	Sequence 254, App
34	60.5	16.6	447	6	US-10-173-692-254	Sequence 254, App
35	60.5	16.6	447	6	US-10-173-695-254	Sequence 254, App
36	60.5	16.6	447	6	US-10-173-697-254	Sequence 254, App
37	60.5	16.6	447	6	US-10-173-698-254	Sequence 254, App
38	60.5	16.6	447	6	US-10-173-699-254	Sequence 254, App
39	60.5	16.6	447	6	US-10-173-701-254	Sequence 254, App
40	60.5	16.6	447	6	US-10-173-704-254	Sequence 254, App
41	60.5	16.6	447	6	US-10-173-708-254	Sequence 254, App
42	60.5	16.6	447	6	US-10-174-571-254	Sequence 254, App
43	60.5	16.6	447	6	US-10-174-574-254	Sequence 254, App
44	60.5	16.6	447	6	US-10-174-583-254	Sequence 254, App
45	60.5	16.6	447	6	US-10-173-694-254	Sequence 254, App

ALIGNMENTS

RESULT 1
US-10-219-999-59588
; Sequence 59588, Application US/10219999
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: cDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)C
; CURRENT APPLICATION NUMBER: US/10/219,999
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/324,109
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/312,544
; PRIOR FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 63520
; SEQ ID NO 59588
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Zea mays
US-10-219-999-59588

Query Match 18.4%; Score 67; DB 6; Length 340;
Best Local Similarity 25.4%; Pred. No. 10;
Matches 18; Conservative 15; Mismatches 22; Indels 16; Gaps 3;

```
Qy      15 RGGCQTSNDLLLEAL-NASSKEEALQ--------IIREKIPEKYLFQFHN-------LNS 58
           | || :        |  : ::|::  |         :::|::|: : |  ||       |||
Db      30 RPGCSSRTPTATSATHHGATKQQVRQGKVSDETKMVQEELPQAWWFDSHNLARPSPWLNS 89

Qy      59 NLDRIFDKTPE 69
              |  : ||| :
Db      90 TLSELDDKTKQ 100
```

RESULT 2
US-09-791-537-132140
; Sequence 132140, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210

```
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 132140
;   LENGTH: 945
;   TYPE: PRT
;   ORGANISM: Drosophila melanogaster
US-09-791-537-132140


  Query Match             18.0%;  Score 65.5;  DB 5;  Length 945;
  Best Local Similarity   32.8%;  Pred. No. 54;
  Matches   21;  Conservative    7;  Mismatches   21;  Indels   15;  Gaps    2;

Qy      10 DGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIP--------------EKYLFQFHN 55
           :|    || | ||:    ||   ||||  :| |||               | :|  |:|
Db     145 NGAGVAGGYQ-SNNYNAAALGMLSKEERAEIQREKAKNPGRNLVKPKWENLEPFLKDFYN 203

Qy      56 LNSN 59
           :: |
Db     204 IHPN 207


RESULT   3
US-10-151-336-3
; Sequence 3, Application US/10151336
; GENERAL INFORMATION:
;  APPLICANT: Mason, Hugh
;  APPLICANT:  Palmer, Kenneth
;  APPLICANT:  Hefferon, Kathleen
;  APPLICANT:  Mor, Tsafrir
;  APPLICANT:  Arntzen, Charles
;  TITLE OF INVENTION: Gemini Virus Vectors for Gene Expression in Plants
;  FILE REFERENCE: 4868/84453
;  CURRENT APPLICATION NUMBER: US/10/151,336
;  CURRENT FILING DATE:  2002-05-20
;  PRIOR APPLICATION NUMBER: US/09/414,276
;  PRIOR FILING DATE: 1999-10-07
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 293
;   TYPE: PRT
;   ORGANISM: bean yellow dwarf virus
US-10-151-336-3


  Query Match             17.9%;  Score 65;  DB 6;  Length 293;
  Best Local Similarity   33.3%;  Pred. No. 15;
  Matches   13;  Conservative   11;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      31 ASSKEEALQIIREKIPEKYLFQFHNLNSNLDRIFDKTPE 69
           |:|||| | :|:|: | ::  :   |  : :::|   ||
Db     133 ATSKEEYLDMIKEEFPHEWATKLQWLEYSANKLFPPQPE 171


RESULT   4
US-10-151-336-6
; Sequence 6, Application US/10151336
; GENERAL INFORMATION:
;  APPLICANT: Mason, Hugh
;  APPLICANT:  Palmer, Kenneth
;  APPLICANT:  Hefferon, Kathleen
;  APPLICANT:  Mor, Tsafrir
;  APPLICANT:  Arntzen, Charles
;  TITLE OF INVENTION: Gemini Virus Vectors for Gene Expression in Plants
;  FILE REFERENCE: 4868/84453
;  CURRENT APPLICATION NUMBER: US/10/151,336
;  CURRENT FILING DATE:  2002-05-20
;  PRIOR APPLICATION NUMBER: US/09/414,276
;  PRIOR FILING DATE: 1999-10-07
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 335
;   TYPE: PRT
;   ORGANISM: bean yellow dwarf virus
US-10-151-336-6


  Query Match             17.9%;  Score 65;  DB 6;  Length 335;
  Best Local Similarity   33.3%;  Pred. No. 17;
  Matches   13;  Conservative   11;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      31 ASSKEEALQIIREKIPEKYLFQFHNLNSNLDRIFDKTPE 69
           |:|||| | :|:|: | ::  :   |  : :::|   ||
Db     133 ATSKEEYLDMIKEEFPHEWATKLQWLEYSANKLFPPQPE 171


RESULT   5
US-09-791-537-33198
; Sequence 33198, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 33198
;   LENGTH: 587
;   TYPE: PRT
;   ORGANISM: Streptococcus sp
US-09-791-537-33198


  Query Match             17.6%;  Score 64;  DB 5;  Length 587;
  Best Local Similarity   28.3%;  Pred. No. 46;
  Matches   17;  Conservative   14;  Mismatches   21;  Indels    8;  Gaps    1;

Qy      12 RSARGGCQTSNDLLLEALNASSKE--------EALQIIREKIPEKYLFQFHNLNSNLDRI 63
           :|| || :|: | : : || :|::          :|||   |   :| |   :| |  : ::
Db     114 KSANGGLKTALDDMADTLNQASRKVNELSAQNKALQAEAEAAAQKALDALNNKNEQIAKL 173


RESULT   6
US-09-893-519A-35
; Sequence 35, Application US/09893519A
; GENERAL INFORMATION:
;  APPLICANT: ANADYS PHARMACEUTICALS, INC.
;  APPLICANT:  THOMPSON, Craig
;  APPLICANT:  MOORE, Jeffrey
;  APPLICANT:  BUURMAN, Ed T.
;  APPLICANT:  BRADLEY, John
;  APPLICANT:  DESILVA, Thamara
;  APPLICANT:  HARRIS, Sandra
;  APPLICANT:  KOMARNITSKY, Svetlana
;  APPLICANT:  MENDILLO, Marc
;  APPLICANT:  MOORE, Daniel
;  APPLICANT:  MCCOY, Melissa
;  APPLICANT:  SANDERSON, Karen
;  APPLICANT:  HAQ, Tariq
;  APPLICANT:  ZHU, Shuhao
;  APPLICANT:  LONG, Fan
;  APPLICANT:  DAVIDOV, Eugene
;  TITLE OF INVENTION: ANTIFUNGAL COMPOUNDS AND METHODS OF USE
;  FILE REFERENCE: 0342/1G548-US2
;  CURRENT APPLICATION NUMBER: US/09/893,519A
;  CURRENT FILING DATE:  2001-06-28
;  PRIOR APPLICATION NUMBER: US 60/215,164
;  PRIOR FILING DATE: 2000-06-29
```

```
;  PRIOR APPLICATION NUMBER: US 60/224,457
;  PRIOR FILING DATE: 2000-08-10
;  NUMBER OF SEQ ID NOS: 146
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
;   LENGTH: 712
;   TYPE: PRT
;   ORGANISM: Candida albicans
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Corresponds to SEQ ID NO: 108
US-09-893-519A-35


  Query Match             17.6%;  Score 64;  DB 5;  Length 712;
  Best Local Similarity   27.8%;  Pred. No. 58;
  Matches   20;  Conservative   11;  Mismatches   17;  Indels   24;  Gaps    2;

Qy      22 NDLLLEALNASSKEEALQIIREKIPEKYL--------FQFHNLNSN---------- 59
           |||    || ||:|: | ||  |:  ::|         :| |: ::
Db      35 NDLQTTLLNTSSEEQNLSIIINKLKMQFLSNNLKNEWVEFQNIVNSLSKFKSLDQICNYL 94

Qy      60 --LDRIFDKTPE 69
             || : |: ||
Db      95 AFLDALRDEKPE 106


RESULT   7
US-10-179-131-6704
; Sequence 6704, Application US/10179131
; GENERAL INFORMATION:
;  APPLICANT: HARE, ROBERTA S.
;  APPLICANT:  SHAW, KAREN J.
;  APPLICANT:  SHIMER JR., GEORGE H.
;  APPLICANT:  KESSLER, MARCO
;  APPLICANT:  NOLLING, JORK
;  APPLICANT:  ZENG, QIANDONG
;  APPLICANT:  GREENE, JONATHAN R.
;  TITLE OF INVENTION: CANDIDA ALBICANS NUCLEIC ACIDS AND POLYPEPTIDES,
;  TITLE OF INVENTION:  AND USES THEREFOR
;  FILE REFERENCE: 2976-4031
;  CURRENT APPLICATION NUMBER: US/10/179,131
;  CURRENT FILING DATE:  2002-06-21
;  NUMBER OF SEQ ID NOS: 10194
; SEQ ID NO 6704
;   LENGTH: 793
;   TYPE: PRT
;   ORGANISM: Candida albicans
US-10-179-131-6704


  Query Match             17.6%;  Score 64;  DB 6;  Length 793;
  Best Local Similarity   27.8%;  Pred. No. 67;
  Matches   20;  Conservative   11;  Mismatches   17;  Indels   24;  Gaps    2;

Qy      22 NDLLLEALNASSKEEALQIIREKIPEKYL--------FQFHNLNSN---------- 59
           |||    || ||:|: | ||  |:  ::|         :| |: ::
Db      43 NDLQTTLLNTSSEEQNLSIIINKLKMQFLSNNLKNEWVEFQNIVNSLSKFKSLDQICNYL 102

Qy      60 --LDRIFDKTPE 69
             || : |: ||
Db     103 AFLDALRDEKPE 114


RESULT   8
US-09-791-537-142650
; Sequence 142650, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 142650
;   LENGTH: 1229
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-09-791-537-142650


  Query Match             17.4%;  Score 63.5;  DB 5;  Length 1229;
  Best Local Similarity   27.4%;  Pred. No. 1.3e+02;
  Matches   17;  Conservative   13;  Mismatches   29;  Indels    3;  Gaps    2;

Qy       4 WGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           ||: : |  |      | | :: || :  |: ||  ::|     |  | : |: |  |  :
Db     776 WGKIETDSSSNNADAQNSAEIALE-VEKSAAEEQKKMIGNL--ENQLTEMHDENEKLMSL 832

Qy      64 FD 65
           ::
Db     833 YE 834


RESULT   9
US-09-791-537-100065
; Sequence 100065, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 100065
;   LENGTH: 343
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-09-791-537-100065


  Query Match             17.3%;  Score 63;  DB 5;  Length 343;
  Best Local Similarity   31.1%;  Pred. No. 31;
  Matches   14;  Conservative    7;  Mismatches   16;  Indels    8;  Gaps    1;

Qy       4 WGEFQVDGRS--------ARGGCQTSNDLLLEALNASSKEEALQI 40
           |||: ||| :          :| |    : |||  |   :| : | :
Db     159 WGEYSVDGLTVKTLPGVFSRDGLDVGSQLLLSTLTPHTKGKVLDV 203


RESULT  10
US-09-791-537-23759
; Sequence 23759, Application US/09791537
; GENERAL INFORMATION:
;  APPLICANT: Bionomix, Inc.
;  APPLICANT:  Debe, Derek
;  APPLICANT:  Danzer, Joseph
;  TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;  TITLE OF INVENTION:  METHODS OF USE THEREOF
;  FILE REFERENCE: 261/210
;  CURRENT APPLICATION NUMBER: US/09/791,537
;  CURRENT FILING DATE:  2001-02-22
;  NUMBER OF SEQ ID NOS: 153055
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23759
;   LENGTH: 1286
;   TYPE: PRT
;   ORGANISM: Saccharomyces cerevisiae
```

US-09-791-537-23759

Query Match 17.2%; Score 62.5; DB 5; Length 1286;
Best Local Similarity 27.4%; Pred. No. 1.9e+02;
Matches 17; Conservative 14; Mismatches 28; Indels 3; Gaps 1;

```
Qy       5 GEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQF---HNLNSNLD 61
           ||| : ||: | |  ::  ::: | |:       : :   :| :   ||   :|:  ||
Db     755 GEFTLAGRAGRRGLDSTGTVIVMAYNSPLSIATFKEVTMGVPTRLQSQFRLTYNMILNLL 814

Qy      62 RI 63
           ||
Db     815 RI 816
```

RESULT 11
PCT-US02-21494-17
; Sequence 17, Application PC/TUS0221494
; GENERAL INFORMATION:
; APPLICANT: BOGOCH, SAMUEL
; APPLICANT: BOGOCH, ELENORE S.
; TITLE OF INVENTION: REPLIKIN PEPTIDES AND METHODS OF USE
; FILE REFERENCE: 09425/469765
; CURRENT APPLICATION NUMBER: PCT/US02/21494
; CURRENT FILING DATE: 2002-07-09
; PRIOR APPLICATION NUMBER: PCT/US02/09240
; PRIOR FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: 60/278,761
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/303,396
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 09/984,056
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: 09/984,057
; PRIOR FILING DATE: 2001-10-26
; NUMBER OF SEQ ID NOS: 711
; SOFTWARE: PatentIn 2.1
; SEQ ID NO 17
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Virus recognin
PCT-US02-21494-17

Query Match 17.0%; Score 62; DB 1; Length 21;
Best Local Similarity 52.4%; Pred. No. 1.3;
Matches 11; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

```
Qy      34 KEEALQIIREKIPEKYLFQFH 54
           |  || |:||| |: :: |||
Db       1 KASALNILREKAPKDFVLQFH 21
```

RESULT 12
US-10-105-232-17
; Sequence 17, Application US/10105232
; GENERAL INFORMATION:
; APPLICANT: BOGOCH, SAMUEL
; APPLICANT: BOGOCH, ELENORE S.
; TITLE OF INVENTION: REPLIKIN PEPTIDES IN RAPID REPLICATION OF GLIOMA CELLS
; TITLE OF INVENTION: AND IN INFLUENZA EPIDEMICS
; FILE REFERENCE: 09425-46904
; CURRENT APPLICATION NUMBER: US/10/105,232
; CURRENT FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: 60/303,396
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/278,761
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 09/146,755
; PRIOR FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: 09/817,144
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 08/198,139
; PRIOR FILING DATE: 1994-02-17
; NUMBER OF SEQ ID NOS: 535
; SOFTWARE: PatentIn 2.1
; SEQ ID NO 17
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Virus recognin
US-10-105-232-17

Query Match 17.0%; Score 62; DB 6; Length 21;
Best Local Similarity 52.4%; Pred. No. 1.3;
Matches 11; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

```
Qy      34 KEEALQIIREKIPEKYLFQFH 54
           |  || |:||| |: :: |||
Db       1 KASALNILREKAPKDFVLQFH 21
```

RESULT 13
US-10-053-853A-1143
; Sequence 1143, Application US/10053853A
; GENERAL INFORMATION:
; APPLICANT: HAYASHI, Hideo
; APPLICANT: SHINAGAWA, Hideo
; APPLICANT: MAKINO, Kozo
; APPLICANT: HAYASHI, Tetsuya
; APPLICANT: OHNISHI, Makoto
; APPLICANT: HATTORI, Masahira
; APPLICANT: KUROKAWA, Ken
; TITLE OF INVENTION: Polynucleotide molecules and polypeptides specific to Enterohe
; TITLE OF INVENTION: Escherichia coli 0157:H7 and use thereof
; FILE REFERENCE: 2002-0060A/WMC/01704
; CURRENT APPLICATION NUMBER: US/10/053,853A
; CURRENT FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: JP2001-112010
; PRIOR FILING DATE: 2001-01-24
; NUMBER OF SEQ ID NOS: 1866
; SEQ ID NO 1143
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Escherichia coli O157:H7
US-10-053-853A-1143

Query Match 17.0%; Score 62; DB 6; Length 329;
Best Local Similarity 35.1%; Pred. No. 39;
Matches 20; Conservative 7; Mismatches 18; Indels 12; Gaps 2;

```
Qy      13 SARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRIFDKTPE 69
           | ||    ||  ||: :: : ||  |: | || |   |||:          |:  ||
Db      11 SFRGKTALSNSTLLQKVSFAGKEYPLEPIDEKTP--ILFQW----------FEARPE 55
```

RESULT 14
US-09-791-537-2671
; Sequence 2671, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 2671
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Schistosoma mansoni
US-09-791-537-2671

Query Match 17.0%; Score 62; DB 5; Length 631;
Best Local Similarity 24.2%; Pred. No. 88;
Matches 16; Conservative 15; Mismatches 19; Indels 16; Gaps 2;

```
Qy       1 TLVWGEFQVDG--------------RSARGGCQTSNDLL-LEALNASSKEEALQIIREK 44
           :| | ::|  |               ||:  |    |::: :|  | :: :: :| :| :
Db     485 SLPWEKYQDSGFDFNLDVEQNQTENSRSSSNGLIVGNEIIPVEPANLNACDDVIQFLRSE 544

Qy      45 IPEKYL 50
           || | |
Db     545 IPGKKL 550
```

RESULT 15
US-09-791-537-113157
; Sequence 113157, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 113157
; LENGTH: 643
; TYPE: PRT
; ORGANISM: Schistosoma mansoni
US-09-791-537-113157

Query Match 17.0%; Score 62; DB 5; Length 643;
Best Local Similarity 24.2%; Pred. No. 91;
Matches 16; Conservative 15; Mismatches 19; Indels 16; Gaps 2;

```
Qy       1 TLVWGEFQVDG--------------RSARGGCQTSNDLL-LEALNASSKEEALQIIREK 44
           :| | ::|  |               ||:  |    |::: :|  | :: :: :| :| :
Db     497 SLPWEKYQDSGFDFNLDVEQNQTENSRSSSNGLIVGNEIIPVEPANLNACDDVIQFLRSE 556

Qy      45 IPEKYL 50
           || | |
Db     557 IPGKKL 562
```

Search completed: October 10, 2002, 02:09:04
Job time: 1447 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:47:19 ; Search time 133.15 Seconds
(without alignments)
58.394 Million cell updates/sec

Title: US-09-289-346A-10
Perfect score: 364
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	364	100.0	70	21	AAB18686	Mutant peptide der
2	349	95.9	70	21	AAB18677	Peptide fragment f
3	349	95.9	356	21	AAB18687	Amino acid sequenc
4	341	93.7	70	21	AAB18685	Mutant peptide der
5	338	92.9	70	21	AAB18688	Mutant peptide der
6	337	92.6	70	21	AAB18692	Mutant peptide der
7	335	92.0	70	21	AAB18684	Mutant peptide der
8	335	92.0	70	21	AAB18690	Mutant peptide der
9	334	91.8	70	21	AAB18678	Mutant peptide der
10	333	91.5	70	21	AAB18689	Mutant peptide der
11	331	90.9	70	21	AAB18680	Mutant peptide der
12	331	90.9	70	21	AAB18691	Mutant peptide der
13	328	90.1	70	21	AAB18681	Mutant peptide der
14	327	89.8	70	21	AAB18683	Mutant peptide der
15	325	89.3	70	21	AAB18682	Mutant peptide der
16	317	87.1	70	21	AAB18679	Mutant peptide der
17	223	61.3	353	18	AAW34338	Bean golden mosaic
18	223	61.3	353	18	AAW34332	Bean golden mosaic
19	223	61.3	353	18	AAW34333	Bean golden mosaic
20	223	61.3	353	18	AAW34334	Bean golden mosaic
21	223	61.3	353	18	AAW34335	Bean golden mosaic
22	221	60.7	353	8	AAP70407	ORF 4 gene product
23	221	60.7	359	17	AAR88870	Sardinian tomato y
24	221	60.7	359	17	AAR88871	Sardinian tomato y
25	221	60.7	359	17	AAR88872	Sardinian tomato y
26	221	60.7	361	18	AAW34336	Tomato mottle viru
27	221	60.7	361	18	AAW34324	Tomato mottle viru
28	221	60.7	361	18	AAW34325	Tomato mottle viru
29	221	60.7	361	18	AAW34326	Tomato mottle viru
30	212.5	58.4	361	8	AAP70562	Product of ORF 4 f
31	207	56.9	362	19	AAW56495	Tobacco leaf curli
32	206	56.6	357	18	AAW34329	Tomato yellow leaf
33	206	56.6	357	18	AAW34330	Tomato yellow leaf
34	206	56.6	357	18	AAW34331	Tomato yellow leaf
35	198	54.4	357	18	AAW34337	Tomato yellow leaf
36	65.5	18.0	945	22	ABB65231	Drosophila melanog
37	65	17.9	292	21	AAY92317	Geminivirus Rep Cl
38	65	17.9	335	21	AAY92318	Mastrevirus Rep in
39	64.5	17.7	665	22	ABB70991	Drosophila melanog
40	63.5	17.4	163	21	AAG60400	Arabidopsis thalia
41	63	17.3	134	16	AAR72393	Agmenellum quadrup
42	63	17.3	343	22	AAU34881	E. coli cellular p
43	62	17.0	108	16	AAR82847	Murine mutant cyto
44	62	17.0	272	22	AAE10442	Novel human phosph
45	62	17.0	340	22	AAE11930	Human CG121 (or C5

ALIGNMENTS

RESULT 1
AAB18686
ID AAB18686 standard; peptide; 70 AA.
XX
AC AAB18686;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 24
FT /note= "wild type residue replaced with Leu"
FT Misc-difference 25
FT /note= "wild type residue replaced with Leu"
FT Misc-difference 26
FT /note= "wild type residue replaced with Leu"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.

```
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 46; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             100.0%;  Score 364;  DB 21;  Length 70;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-38;
  Matches   70;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 tlvwgefqvdgrsarggcqtsndlllealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   2
AAB18677
ID   AAB18677 standard; peptide; 70 AA.
XX
AC   AAB18677;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Peptide fragment from Rep protein of TGMV (amino acids 110-179).
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Tomato golden mosaic virus.
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 18; 73pp; English.
XX
CC   The present sequence is derived from a geminivirus replication (Rep)
CC   protein, which is also known as AL1. AL1 binds double-stranded DNA,
CC   catalyses cleavage and ligation of single-stranded DNA, and interacts
CC   with other viral and host proteins. Mutants of the AL1 protein are used
CC   to produce transgenic plants. The mutation in AL1 is present in a
CC   ribosome binding region, and expression of mutant AL1 protein imparts
CC   increased resistance to geminivirus infection in the plant. Mutant AL1
CC   proteins are useful for producing plants having increased resistance or
CC   reduced sensitivity to a geminivirus such as tomato golden mosaic virus,
CC   tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl
CC   virus, African cassava mosaic virus, Indian cassava mosaic virus, potato
CC   yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,
CC   squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or
CC   beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             95.9%;  Score 349;  DB 21;  Length 70;
  Best Local Similarity   95.7%;  Pred. No. 9.2e-37;
  Matches   67;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          ||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db      1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   3
AAB18687
ID   AAB18687 standard; peptide; 356 AA.
XX
AC   AAB18687;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Amino acid sequence of a geminivirus replication protein of TGMV.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 354
FT                   /note= "unspecified amino acid"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
```

PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT -
XX
PS Disclosure; Page 47-48; 73pp; English.
XX
CC The present sequence represents a geminivirus replication (Rep)
CC protein, which is also known as AL1. AL1 binds double-stranded DNA,
CC catalyses cleavage and ligation of single-stranded DNA, and interacts
CC with other viral and host proteins. Mutants of the AL1 protein are used
CC to produce transgenic plants. The mutation in AL1 is present in a
CC ribosome binding region, and expression of mutant AL1 protein imparts
CC increased resistance to geminivirus infection in the plant. Mutant AL1
CC proteins are useful for producing plants having increased resistance or
CC reduced sensitivity to a geminivirus such as tomato golden mosaic virus,
CC tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl
CC virus, African cassava mosaic virus, Indian cassava mosaic virus, potato
CC yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,
CC squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or
CC beet curly top virus.
XX
SQ Sequence 356 AA;

```
  Query Match             95.9%;  Score 349;  DB 21;  Length 356;
  Best Local Similarity   95.7%;  Pred. No. 7.2e-36;
  Matches   67;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||
Db     110 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 169

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     170 drifdktpep 179
```

RESULT 4
AAB18685
ID AAB18685 standard; peptide; 70 AA.
XX
AC AAB18685;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 10
FT /note= "wild type residue replaced with Ala"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DR WPI; 2000-618851/59.
XX
PT Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT -
XX
PS Claim 53; Page 46; 73pp; English.
XX
CC The present sequence represents a mutant peptide, derived from a
CC geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC protein are used to produce transgenic plants. The mutation in AL1 is
CC present in a ribosome binding region, and expression of mutant AL1
CC protein imparts increased resistance to geminivirus infection in the
CC plant. Mutant AL1 proteins are useful for producing plants having
CC increased resistance or reduced sensitivity to a geminivirus such as
CC tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC virus, cotton leaf curl virus or beet curly top virus.
XX
SQ Sequence 70 AA;

```
  Query Match             93.7%;  Score 341;  DB 21;  Length 70;
  Best Local Similarity   94.3%;  Pred. No. 9.6e-36;
  Matches   66;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||| ||||||||||||   ||||||||||||||||||||||||||||||||
Db       1 tlvwgefqvagrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70
```

RESULT 5
AAB18688
ID AAB18688 standard; peptide; 70 AA.
XX
AC AAB18688;
XX
DT 22-JAN-2001 (first entry)
XX
DE Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Synthetic.
OS Tomato golden mosaic virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 19
FT /note= "wild type residue replaced with Ala"
FT Misc-difference 20
FT /note= "wild type residue replaced with Ala"
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DR WPI; 2000-618851/59.

```
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 48; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             92.9%;  Score 338;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 2.3e-35;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |||||||||||||||||  |||   ||||||||||||||||||||||||||||||||||||
Db      1 tlvwgefqvdgrsarggcaasndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   6
AAB18692
ID   AAB18692 standard; peptide; 70 AA.
XX
AC   AAB18692;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 66
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 69
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 50; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             92.6%;  Score 337;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 3.1e-35;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db      1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          |||||  || |
Db     61 drifdatpap 70


RESULT   7
AAB18684
ID   AAB18684 standard; peptide; 70 AA.
XX
AC   AAB18684;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 7
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 8
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
```

```
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             92.0%;  Score 335;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 5.5e-35;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          ||||||  ||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db      1 tlvwgeaavdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   8
AAB18690
ID   AAB18690 standard; peptide; 70 AA.
XX
AC   AAB18690;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 27
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 30
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             92.0%;  Score 335;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 5.5e-35;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          ||||||||||||||||||||||    || ||||||||||||||||||||||||||||||
Db      1 tlvwgefqvdgrsarggcqtsndaaaaalaasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   9
AAB18678
ID   AAB18678 standard; peptide; 70 AA.
XX
AC   AAB18678;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 12
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 13
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 15
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
```

```
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 42-43; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             91.8%;  Score 334;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 7.4e-35;
  Matches   64;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||  :| ||||||||   |||||||||||||||||||||||||||||||||
Db       1 tlvwgefqvdgaaaaggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  10
AAB18689
ID   AAB18689 standard; peptide; 70 AA.
XX
AC   AAB18689;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 22
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 23
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 48-49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             91.5%;  Score 333;  DB 21;  Length 70;
  Best Local Similarity   92.9%;  Pred. No. 9.9e-35;
  Matches   65;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||||||||||||     |||||||||||||||||||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsaaaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  11
AAB18680
ID   AAB18680 standard; peptide; 70 AA.
XX
AC   AAB18680;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 42
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 43
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 44
```

```
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 43-44; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             90.9%;  Score 331;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 1.8e-34;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||   ||||||||||||||   ||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiiaaaipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  12
AAB18691
ID   AAB18691 standard; peptide; 70 AA.
XX
AC   AAB18691;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 34
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 35
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 36
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             90.9%;  Score 331;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 1.8e-34;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||   |||||||   ||||||||||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnassaaaalqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  13
AAB18681
ID   AAB18681 standard; peptide; 70 AA.
XX
AC   AAB18681;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
```

```
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 47
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 48
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 49
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 44; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             90.1%;  Score 328;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 4.3e-34;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||   ||||||||||||||||||||   ||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipaaalfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  14
AAB18683
ID   AAB18683 standard; peptide; 70 AA.
XX
AC   AAB18683;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 59
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 61
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 62
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             89.8%;  Score 327;  DB 21;  Length 70;
  Best Local Similarity   91.4%;  Pred. No. 5.8e-34;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||   |||||||||||||||||||||||||||||| |
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsal 60

Qy      61 DRIFDKTPEP 70
           ||||||||
Db      61 aaifdktpep 70


RESULT  15
AAB18682
ID   AAB18682 standard; peptide; 70 AA.
XX
AC   AAB18682;
XX
```

```
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 52
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 54
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 55
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 44-45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match           89.3%;  Score 325;  DB 21;  Length 70;
  Best Local Similarity  91.4%;  Pred. No. 1e-33;
  Matches   64;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          ||||||||||||||||||||||   |||||||||||||||||||||||| |  ||||||
Db      1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfafaalnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


Search completed: October 10, 2002, 01:47:19

Job time: 528 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:09:49 ; Search time 34.62 Seconds
(without alignments)
78.289 Million cell updates/sec

Title: US-09-289-346A-2
Perfect score: 362
Sequence: 1 TLVWGEFQVDGAAAAGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	349	96.4	352	1	VAL1_TGMV	P03567 tomato gold
2	249	68.8	361	1	VAL1_PYMVV	P27258 potato yell
3	236	65.2	358	1	VAL1_CLVK	P14982 cassava lat
4	236	65.2	358	1	VAL1_CLVN	P14972 cassava lat
5	229	63.3	362	1	VAL1_TYLCA	P36279 tomato yell
6	222	61.3	359	1	VAL1_TYLCU	P38609 tomato yell
7	218	60.2	349	1	VAL1_PHUV	Q06923 pepper huas
8	215	59.4	359	1	VAL1_TYLCM	P27260 tomato yell
9	213	58.8	353	1	VAL1_BGMV	P05175 bean golden
10	212	58.6	355	1	VAL1_ABMVW	P21947 abutilon mo
11	209	57.7	358	1	VAL1_BCTV	P14991 beet curly
12	209	57.7	361	1	VAL1_TMOV	Q06657 tomato mott
13	200	55.2	357	1	VAL1_TYLCV	P27259 tomato yell
14	124	34.3	347	1	VAL1_SLCV	P29048 squash leaf
15	68.5	18.9	1713	1	LMA3_HUMAN	Q16787 homo sapien
16	65.5	18.1	1610	1	CCAD_MESAU	Q99244 mesocricetu
17	65.5	18.1	2161	1	CCAD_HUMAN	Q01668 homo sapien
18	65.5	18.1	2203	1	CCAD_RAT	P27732 rattus norv
19	64.5	17.8	299	1	Y175_HELPJ	Q9zmq7 helicobacte
20	64.5	17.8	334	1	G3P_BACST	P00362 bacillus st
21	64.5	17.8	335	1	G3P_BACCO	P15115 bacillus co
22	64.5	17.8	557	1	SASB_ANAPL	Q04791 anas platyr
23	62.5	17.3	419	1	Y223_MYCPN	P75465 mycoplasma
24	62	17.1	295	1	VAL1_TYDVA	P31617 tobacco yel
25	62	17.1	1483	1	CYP1_YEAST	P12351 saccharomyc
26	60.5	16.7	136	1	Y452_CAEEL	O62250 caenorhabdi
27	60	16.6	387	1	VE2_HPV41	P27552 human papil
28	60	16.6	447	1	CDSA_DROME	P56079 d phosphati
29	60	16.6	513	1	HEM0_CHICK	P18080 gallus gall
30	59.5	16.4	863	1	AMPN_CAUCR	P37893 caulobacter
31	59	16.3	129	1	RK12_PORPU	P51339 porphyra pu
32	59	16.3	247	1	YCP4_YEAST	P25349 saccharomyc
33	58.5	16.2	299	1	Y175_HELPY	P56112 helicobacte
34	58.5	16.2	439	1	XYLA_LACLA	Q9cfg7 lactococcus
35	58.5	16.2	589	1	YE67_HAEIN	P45221 haemophilus
36	58.5	16.2	1044	1	BUB1_SCHPO	O94751 schizosacch
37	58.5	16.2	2190	1	CCAD_CHICK	O73700 gallus gall
38	58	16.0	387	1	Y4PF_RHISN	P55615 rhizobium s
39	58	16.0	422	1	BCHN_ACIRU	Q9wxb4 acidiphiliu
40	58	16.0	617	1	YACH_ECOLI	P36682 escherichia
41	58	16.0	640	1	DNAK_RHOMR	Q9xcb1 rhodothermu
42	58	16.0	874	1	SLAP_BACLI	P49052 bacillus li
43	57.5	15.9	266	1	ETFB_MYCLE	O33095 mycobacteri
44	57.5	15.9	266	1	ETFB_MYCTU	O53276 mycobacteri
45	57.5	15.9	525	1	P2BB_RAT	P20651 rattus norv

ALIGNMENTS

RESULT 1
VAL1_TGMV
ID VAL1_TGMV STANDARD; PRT; 352 AA.
AC P03567;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein.
GN AC1.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10831;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;
RT "Complete nucleotide sequence of the infectious cloned DNA components
RT of tomato golden mosaic virus: potential coding regions and regulatory
RT sequences.";
RL EMBO J. 3:2197-2205(1984).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; K02029; -; NOT_ANNOTATED_CDS.
DR PIR; A04170; QQCVL1.
DR InterPro; IPR001191; Gemini_AL1.
DR Pfam; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR ProDom; PD000736; Gemini_AL1; 1.
KW ATP-binding.
FT NP_BIND 223 230 ATP (POTENTIAL).
SQ SEQUENCE 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;

Query Match 96.4%; Score 349; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 5e-31;
Matches 67; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||| :| ||||||||||||||||||||||||||||||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
VAL1_PYMVV
ID VAL1_PYMVV STANDARD; PRT; 361 AA.

```
AC   P27258;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 protein.
OS   Potato yellow mosaic virus (isolate Venezuela).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10828;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91311403; PubMed=1856690;
RA   Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT   "The nucleotide sequence of the infectious cloned DNA components of
RT   potato yellow mosaic virus.";
RL   J. Gen. Virol. 72:1515-1520(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; D00940; BAA00782.1; -.
DR   PIR; JU0364; QQCVPT.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   361 AA;  40850 MW;  5627A33BF1264383 CRC64;


  Query Match             68.8%;  Score 249;  DB 1;  Length 361;
  Best Local Similarity   66.7%;  Pred. No. 4.7e-20;
  Matches   46;  Conservative   12;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:|| :| || || |||||||||: :|| |::|||:||:|||:|||:|||: ||
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 DRIFDKTPE 69
           |||| | ||
Db     170 DRIFMKAPE 178


RESULT   3
VAL1_CLVK
ID   VAL1_CLVK      STANDARD;      PRT;   358 AA.
AC   P14982;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (40.4 kDa protein).
GN   AC1.
OS   Cassava latent virus (strain West Kenyan 844).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10818;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Gay M.R.;
RT   "Nucleotide sequence of cassava latent virus DNA.";
RL   Nature 301:260-262(1983).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; J02057; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40346 MW;  ED173E753EE92D69 CRC64;


  Query Match             65.2%;  Score 236;  DB 1;  Length 358;
  Best Local Similarity   60.0%;  Pred. No. 1.3e-18;
  Matches   42;  Conservative   14;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|| :| || |::||| |:|||: || ||| :||| :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   4
VAL1_CLVN
ID   VAL1_CLVN      STANDARD;      PRT;   358 AA.
AC   P14972;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (40.4 kDa protein).
GN   AC1.
OS   Cassava latent virus (strain Nigerian).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10819;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90174930; PubMed=2308831;
RA   Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT   "Nucleotide sequence of the infectious cloned DNA components of
RT   African cassava mosaic virus (Nigerian strain).";
RL   Nucleic Acids Res. 18:197-198(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X17095; CAA34953.1; -.
DR   PIR; S07594; S07594.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40435 MW;  1DB16BB0CB2D5E2C CRC64;


  Query Match             65.2%;  Score 236;  DB 1;  Length 358;
  Best Local Similarity   60.0%;  Pred. No. 1.3e-18;
  Matches   42;  Conservative   14;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|| :| || |::||| |:|||: || ||| :||| :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168
```

```
Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   5
VAL1_TYLCA
ID   VAL1_TYLCA     STANDARD;      PRT;   362 AA.
AC   P36279;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=36447;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93139778; PubMed=8423446;
RA   Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;
RT   "Nucleotide sequence and genome organization of tomato leaf curl
RT   geminivirus.";
RL   J. Gen. Virol. 74:147-151(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
DR   PIR; JQ1887; JQ1887.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   362 AA;  41197 MW;  343E7184B4704098 CRC64;


  Query Match             63.3%;  Score 229;  DB 1;  Length 362;
  Best Local Similarity   52.9%;  Pred. No. 7.4e-18;
  Matches   45;  Conservative   10;  Mismatches   14;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || ||||:|| :| || |::||| |:||| || ||| ::|| |: |: |||||||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   6
VAL1_TYLCU
ID   VAL1_TYLCU     STANDARD;      PRT;   359 AA.
AC   P38609;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=37139;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94256836; PubMed=8198442;
RA   Noris E., Hidalgo E., Accotto G., Moriones E.;
RT   "High similarity among the tomato yellow leaf curl virus isolates
RT   from the west Mediterranean basin: the nucleotide sequence of an
RT   infectious clone from Spain.";
RL   Arch. Virol. 135:165-170(1994).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z25751; CAA81026.1; -.
DR   PIR; S39211; S39211.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  41065 MW;  2D170A51EF80A3EC CRC64;


  Query Match             61.3%;  Score 222;  DB 1;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 4.3e-17;
  Matches   40;  Conservative   10;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|| :| || ||:||| |:|:|| || ||| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   7
VAL1_PHUV
ID   VAL1_PHUV      STANDARD;      PRT;   349 AA.
AC   Q06923;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AL1.
OS   Pepper huasteco virus (PHV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=28349;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94015007; PubMed=8409944;
RA   Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA   Rivera-Bustamante R.F.;
RT   "Complete nucleotide sequence of pepper huasteco virus: analysis and
RT   comparison with bipartite geminiviruses.";
RL   J. Gen. Virol. 74:2225-2231(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70418; CAA49856.1; -.
DR   PIR; S31875; S31875.
DR   PIR; JQ2300; JQ2300.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (BY SIMILARITY).
SQ   SEQUENCE   349 AA;  39722 MW;  D5F4E76CD56370F4 CRC64;


  Query Match             60.2%;  Score 218;  DB 1;  Length 349;
```

```
  Best Local Similarity   57.1%;  Pred. No. 1.1e-16;
  Matches   40;  Conservative   14;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|| :| || |::||  |:|||::| |||||||:|: |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   8
VAL1_TYLCM
ID   VAL1_TYLCM     STANDARD;      PRT;   359 AA.
AC   P27260;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10833;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92107660; PubMed=1840676;
RA   Kheyr-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,
RA   Gronenborn B.;
RT   "Tomato yellow leaf curl virus from Sardinia is a
RT   whitefly-transmitted monopartite geminivirus.";
RL   Nucleic Acids Res. 19:6763-6769(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X61153; CAA43466.1; -.
DR   PIR; S22593; S22593.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  40733 MW;  9717B4A07C93EFA7 CRC64;


  Query Match             59.4%;  Score 215;  DB 1;  Length 359;
  Best Local Similarity   55.1%;  Pred. No. 2.5e-16;
  Matches   38;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|| :| || ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   9
VAL1_BGMV
ID   VAL1_BGMV      STANDARD;      PRT;   353 AA.
AC   P05175;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   AL1 protein (40.2 kDa protein).
GN   AC1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Howarth A.J., Caton J., Bossert M., Goodman R.M.;
RT   "Nucleotide sequence of bean golden mosaic virus and a model for gene
RT   regulation in geminiviruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M10070; AAA46318.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   353 AA;  40190 MW;  80FA779DF6029A34 CRC64;


  Query Match             58.8%;  Score 213;  DB 1;  Length 353;
  Best Local Similarity   57.1%;  Pred. No. 4.1e-16;
  Matches   40;  Conservative   14;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||||| :| || |::||: |:|||| | | || |::|: |: |: | ||: |||
Db     110 TIEWGQFQVDGRSARGGQQSANDSYAKALNADSIESALTILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFFKVPEP 179


RESULT  10
VAL1_ABMVW
ID   VAL1_ABMVW     STANDARD;      PRT;   355 AA.
AC   P21947;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AC1.
OS   Abutilon mosaic virus (isolate West India).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10816;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91020984; PubMed=2219703;
RA   Frischmuth T., Zimmat G., Jeske H.;
RT   "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT   as well as eukaryotic features.";
RL   Virology 178:461-468(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR   PIR; A36214; QQCVW1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   355 AA;  40257 MW;  16A2CA8A63251E95 CRC64;


  Query Match             58.6%;  Score 212;  DB 1;  Length 355;
  Best Local Similarity   55.7%;  Pred. No. 5.3e-16;
  Matches   39;  Conservative   13;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |  ||||:|| :| || ||:||: |:||||   : || |::|: |: |: | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
VAL1_BCTV
ID   VAL1_BCTV      STANDARD;      PRT;   358 AA.
AC   P14991;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 protein (40.8 kDa protein).
OS   Beet curly top virus (BCTV).
OC   Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
OX   NCBI_TaxID=10840;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT   "The nucleotide sequence of an infectious clone of the geminivirus
RT   beet curly top virus.";
RL   EMBO J. 5:1761-1767(1986).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40889 MW;  39A45FE3C0B9C333 CRC64;


  Query Match             57.7%;  Score 209;  DB 1;  Length 358;
  Best Local Similarity   52.9%;  Pred. No. 1.1e-15;
  Matches   37;  Conservative   17;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|| :| || ||:||: |:|||||:| ::||||::|: |: |  | ||| :|
Db     110 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 169

Qy      61 DRIFDKTPEP 70
           :|| : |:|
Db     170 QKIFQRPPDP 179


RESULT  12
VAL1_TMOV
ID   VAL1_TMOV      STANDARD;      PRT;   361 AA.
AC   Q06657;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein.
GN   AL1.
OS   Tomato mottle virus (isolate Florida) (TMoV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=36449;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93107858; PubMed=1469361;
RA   Abouzid A.M., Polston J.E., Hiebert E.;
RT   "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT   isolated from tomatoes in Florida.";
RL   J. Gen. Virol. 73:3225-3229(1992).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L14460; AAC32414.1; -.
DR   PIR; JQ1870; JQ1870.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (BY SIMILARITY).
SQ   SEQUENCE   361 AA;  40516 MW;  8138B65CEEAC6950 CRC64;


  Query Match             57.7%;  Score 209;  DB 1;  Length 361;
  Best Local Similarity   52.9%;  Pred. No. 1.2e-15;
  Matches   37;  Conservative   17;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|| :| || |::||: |:|||| | : || ::||: |: :: | ||: |||
Db     110 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  13
VAL1_TYLCV
ID   VAL1_TYLCV     STANDARD;      PRT;   357 AA.
AC   P27259;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 protein (C1 protein).
GN   C1.
OS   Tomato yellow leaf curl virus (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10832;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92024070; PubMed=1926771;
RA   Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;
RT   "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RT   with a single genomic component.";
RL   Virology 185:151-161(1991).
```

```
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15656; CAA33688.1; -.
DR   PIR; D40779; QQCVC1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     219    226       ATP (POTENTIAL).
SQ   SEQUENCE   357 AA;  40678 MW;  939AB68E1AB3B2A7 CRC64;


  Query Match             55.2%;  Score 200;  DB 1;  Length 357;
  Best Local Similarity   62.3%;  Pred. No. 1.1e-14;
  Matches   38;  Conservative   11;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           :|  |:|| :| || |::||| |||||: || ||| |::|| |: |: |||||:||||||
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171


RESULT  14
VAL1_SLCV
ID   VAL1_SLCV      STANDARD;      PRT;   347 AA.
AC   P29048;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   AL1 protein.
OS   Squash leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10829;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91082449; PubMed=1984668;
RA   Lazarowitz S.G., Lazdins I.B.;
RT   "Infectivity and complete nucleotide sequence of the cloned genomic
RT   components of a bipartite squash leaf curl geminivirus with a broad
RT   host range phenotype.";
RL   Virology 180:58-69(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M38183; AAC32410.1; ALT_INIT.
DR   PIR; C36785; QQCVS1.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
KW   ATP-binding.
FT   NP_BIND     218    225       ATP (POTENTIAL).
SQ   SEQUENCE   347 AA;  39110 MW;  AFDAEBDDE122110E CRC64;


  Query Match             34.3%;  Score 124;  DB 1;  Length 347;
  Best Local Similarity   37.9%;  Pred. No. 2.3e-06;
  Matches   25;  Conservative   14;  Mismatches   23;  Indels    4;  Gaps    1;

Qy       5 GEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRIF 64
           |:::| |    |     :|    |:|| |  ||| ||:   |: ::  :||| :|::|:|
Db     116 GQYKVSG----GSKSNKDDVYHNAVNAGSAGEALDIIKAGDPKTFIVNYHNLLANVERLF 171

Qy      65 DKTPEP 70
             | |||
Db     172 QKPPEP 177


RESULT  15
LMA3_HUMAN
ID   LMA3_HUMAN     STANDARD;      PRT;  1713 AA.
AC   Q16787; Q13679; Q13680;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   Laminin alpha-3 chain precursor (Epiligrin 170 kDa subunit) (E170)
DE   (Nicein alpha subunit).
GN   LAMA3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Keratinocytes;
RX   MEDLINE=94357926; PubMed=8077230;
RA   Ryan M.C., Tizard R., Vandevanter D.R., Carter W.G.;
RT   "Cloning of the LamA3 gene encoding the alpha 3 chain of the adhesive
RT   ligand epiligrin. Expression in wound repair.";
RL   J. Biol. Chem. 269:22779-22787(1994).
RN   [2]
RP   SEQUENCE OF 1-1331 FROM N.A. (ISOFORMS A AND B).
RX   MEDLINE=96163880; PubMed=8586427;
RA   Vidal F., Baudoin C., Miquel C., Galliano M.-F., Christiano A.M.,
RA   Uitto J., Ortonne J.-P., Meneguzzi G.;
RT   "Cloning of the laminin alpha 3 chain gene (LAMA3) and identification
RT   of a homozygous deletion in a patient with Herlitz junctional
RT   epidermolysis bullosa.";
RL   Genomics 30:273-280(1995).
CC   -!- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
CC       IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
CC       CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
CC       WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC   -!- FUNCTION: LAMININ-5 IS THOUGHT TO BE INVOLVED IN (1) CELL ADHESION
CC       VIA INTEGRIN ALPHA-3/BETA-1 IN FOCAL ADHESION AND INTEGRIN ALPHA-
CC       6/BETA-4 IN HEMIDESMOSOMES, (2) SIGNAL TRANSDUCTION VIA TYROSINE
CC       PHOSPHORYLATION OF PP125-FAK AND P80, (3) DIFFERENTIATION OF
CC       KERATINOCYTES.
CC   -!- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
CC       DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
CC       TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
CC       COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC       THE ALPHA-3 CHAIN IS A SUBUNIT OF LAMININ-5 (EPILIGRIN/KALININ/
CC       NICEIN), AND POSSIBLY ALSO A COMPONENT OF LAMININ-6 (K-LAMININ)
CC       AND LAMININ-7 (KS-LAMININ).
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
CC       MEMBRANES (MAJOR COMPONENT).
CC   -!- ALTERNATIVE PRODUCTS: THE TWO ISOFORMS A AND B DIFFER IN THEIR N-
CC       TERMINUS. THE SEQUENCE SHOWN HERE IS THAT OF THE SMALLER VARIANT
CC       A.
CC   -!- TISSUE SPECIFICITY: SKIN; RESPIRATORY, URINARY, AND DIGESTIVE
CC       EPITHELIA AND IN OTHER SPECIALIZED TISSUES WITH PROMINENT
CC       SECRETORY OR PROTECTIVE FUNCTIONS. EPITHELIAL BASEMENT MEMBRANE,
CC       AND EPITHELIAL CELL TONGUE THAT MIGRATES INTO A WOUND BED. A
CC       DIFFERENTIAL AND FOCAL EXPRESSION OF THE ALPHA-3 CHAIN IS OBSERVED
CC       IN THE CNS.
CC   -!- INDUCTION: LAMININ-5 IS UP-REGULATED IN WOUND SITES OF HUMAN SKIN.
```

```
CC   -!- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC       WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC   -!- DOMAIN: DOMAIN G IS GLOBULAR.
CC   -!- DISEASE: DEFECTS IN LAMA3 ARE A CAUSE OF JUNCTIONAL EPIDERMOLYSIS
CC       BULLOSA (JEB) GRAVIS, ALSO KNOWN AS HERLITZ TYPE JEB. IT IS A
CC       BLISTERING DISORDER IN SKIN THAT IS CHARACTERIZED BY A SEPARATION
CC       OF BASAL CELLS FROM THE BASEMENT MEMBRANE DUE TO A DECREASED
CC       NUMBER OF HEMIDESMOSOMES. LAMININ-5 IS MISSING FROM THE BASEMENT
CC       MEMBRANE OF PATIENTS WITH THE GRAVIS FORM OF EPIDERMOLYSIS
CC       BULLOSA.
CC   -!- SIMILARITY: CONTAINS 2.5 LAMININ EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L34155; AAA59483.1; -.
DR   EMBL; X85107; CAA59428.1; -.
DR   EMBL; X85108; CAA59429.1; -.
DR   HSSP; P02468; 1TLE.
DR   MIM; 600805; -.
DR   MIM; 226700; -.
DR   InterPro; IPR000561; EGF-like.
DR   InterPro; IPR002049; Laminin_EGF.
DR   InterPro; IPR001791; Laminin_G.
DR   InterPro; IPR003129; TSPN.
DR   Pfam; PF00053; laminin_EGF; 2.
DR   Pfam; PF00054; laminin_G; 2.
DR   SMART; SM00180; EGF_Lam; 2.
DR   SMART; SM00282; LamG; 5.
DR   SMART; SM00210; TSPN; 1.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01248; LAMININ_TYPE_EGF; 2.
DR   PROSITE; PS50025; LAM_G_DOMAIN; 5.
KW   Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW   Laminin EGF-like domain; Cell adhesion; Repeat; Signal;
KW   Alternative splicing; Epidermolysis bullosa.
FT   SIGNAL        1     20       POTENTIAL.
FT   CHAIN        21   1713       LAMININ ALPHA-3 CHAIN.
FT   DOMAIN       46    201       DOMAIN III A.
FT   DOMAIN       67    113       LAMININ EGF-LIKE 1.
FT   DOMAIN      114    166       LAMININ EGF-LIKE 2.
FT   DOMAIN      167    185       LAMININ EGF-LIKE 3 (INCOMPLETE).
FT   DOMAIN      186    769       DOMAIN II AND I.
FT   DOMAIN      770    971       LAMININ G-LIKE 1.
FT   DOMAIN      978   1140       LAMININ G-LIKE 2.
FT   DOMAIN     1147   1307       LAMININ G-LIKE 3.
FT   DOMAIN     1366   1530       LAMININ G-LIKE 4.
FT   DOMAIN     1537   1710       LAMININ G-LIKE 5.
FT   DOMAIN      231    327       COILED COIL (POTENTIAL).
FT   DOMAIN      396    548       COILED COIL (POTENTIAL).
FT   DOMAIN      594    621       COILED COIL (POTENTIAL).
FT   DOMAIN      702    765       COILED COIL (POTENTIAL).
FT   DOMAIN     1686   1713       COILED COIL (POTENTIAL).
FT   DISULFID     67     76       BY SIMILARITY.
FT   DISULFID     69     83       BY SIMILARITY.
FT   DISULFID     86     95       BY SIMILARITY.
FT   DISULFID     98    111       BY SIMILARITY.
FT   DISULFID    114    126       BY SIMILARITY.
FT   DISULFID    116    135       BY SIMILARITY.
FT   DISULFID    137    146       BY SIMILARITY.
FT   DISULFID    149    164       BY SIMILARITY.
FT   DISULFID    202    202       INTERCHAIN (PROBABLE).
FT   DISULFID    205    205       INTERCHAIN (PROBABLE).
FT   SITE        658    660       CELL ATTACHMENT SITE (POTENTIAL).
FT   CARBOHYD    542    542       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    645    645       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    745    745       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    882    882       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    964    964       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1108   1108       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1131   1131       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1325   1325       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1477   1477       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1667   1667       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC      1     45       MGWLWIFGAALGQCLGYSSQQQRVPFLQPPGQSQLQASYVE
FT                                FRPS -> KVSSYGGYLTYQAKSFALPGDMVLLEKKPDVQL
FT                                TGQHMSIIYEETNTPRPDRLHHGRVHVVEGNFRHASSRAPV
FT                                SREELMTVLSRLADVRIQGLYFTETQRLTLSEVGLEEASDT
FT                                GSGRIALAVEICACPPAYAGDSC (IN ISOFORM B).
FT   CONFLICT      5      5       W -> R (IN REF. 2).
FT   CONFLICT    123    125       ATG -> GMC (IN REF. 2).
FT   CONFLICT    481    481       M -> K (IN REF. 2).
FT   CONFLICT    754    754       R -> L (IN REF. 2).
FT   CONFLICT    969    969       E -> Q (IN REF. 2).
FT   CONFLICT   1052   1052       D -> A (IN REF. 2).
FT   CONFLICT   1184   1184       G -> A (IN REF. 2).
SQ   SEQUENCE   1713 AA;  189304 MW;  45EA9BE1017B60D3 CRC64;
```

```
  Query Match             18.9%;  Score 68.5;  DB 1;  Length 1713;
  Best Local Similarity   36.1%;  Pred. No. 16;
  Matches   22;  Conservative   11;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       9 VDGAAAAGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPEKYLFQFHNLNSNLDRIFD 65
           || | |      :  || :| |  ||: | ||| :|:| :| |       |:|| |:: :
Db     553 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLLN 608

Qy      66 K 66
           :
Db     609 E 609


Search completed: October 10, 2002, 02:09:50
Job time: 1348 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 01:38:31 ; Search time 133.15 Seconds
(without alignments)
58.394 Million cell updates/sec

Title: US-09-289-346A-1
Perfect score: 364
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	364	100.0	70	21	AAB18677	Peptide fragment f
2	364	100.0	356	21	AAB18687	Amino acid sequenc
3	356	97.8	70	21	AAB18685	Mutant peptide der
4	353	97.0	70	21	AAB18688	Mutant peptide der
5	352	96.7	70	21	AAB18692	Mutant peptide der
6	350	96.2	70	21	AAB18684	Mutant peptide der
7	350	96.2	70	21	AAB18690	Mutant peptide der
8	349	95.9	70	21	AAB18678	Mutant peptide der
9	349	95.9	70	21	AAB18686	Mutant peptide der
10	348	95.6	70	21	AAB18689	Mutant peptide der
11	346	95.1	70	21	AAB18680	Mutant peptide der
12	346	95.1	70	21	AAB18691	Mutant peptide der
13	343	94.2	70	21	AAB18681	Mutant peptide der
14	342	94.0	70	21	AAB18683	Mutant peptide der
15	340	93.4	70	21	AAB18682	Mutant peptide der
16	332	91.2	70	21	AAB18679	Mutant peptide der
17	230	63.2	353	18	AAW34338	Bean golden mosaic
18	230	63.2	353	18	AAW34332	Bean golden mosaic
19	230	63.2	353	18	AAW34333	Bean golden mosaic
20	230	63.2	353	18	AAW34334	Bean golden mosaic
21	230	63.2	353	18	AAW34335	Bean golden mosaic
22	230	63.2	359	17	AAR88870	Sardinian tomato y
23	230	63.2	359	17	AAR88871	Sardinian tomato y
24	230	63.2	359	17	AAR88872	Sardinian tomato y
25	228	62.6	353	8	AAP70407	ORF 4 gene product
26	228	62.6	361	18	AAW34336	Tomato mottle viru
27	228	62.6	361	18	AAW34324	Tomato mottle viru
28	228	62.6	361	18	AAW34325	Tomato mottle viru
29	228	62.6	361	18	AAW34326	Tomato mottle viru
30	221.5	60.9	361	8	AAP70562	Product of ORF 4 f
31	216	59.3	362	19	AAW56495	Tobacco leaf curli
32	215	59.1	357	18	AAW34329	Tomato yellow leaf
33	215	59.1	357	18	AAW34330	Tomato yellow leaf
34	215	59.1	357	18	AAW34331	Tomato yellow leaf
35	207	56.9	357	18	AAW34337	Tomato yellow leaf
36	71.5	19.6	945	22	ABB65231	Drosophila melanog
37	67.5	18.5	512	19	AAW68473	HIV-1 strain YBF30
38	66.5	18.3	665	22	ABB70991	Drosophila melanog
39	65	17.9	292	21	AAY92317	Geminivirus Rep Cl
40	65	17.9	335	21	AAY92318	Mastrevirus Rep in
41	64.5	17.7	299	18	AAW55450	H. pylori ORF 02ae
42	64.5	17.7	1693	21	AAB48457	Human laminin 5 po
43	64.5	17.7	1693	21	AAB48459	Human laminin 5 po
44	64.5	17.7	1713	16	AAR70148	Deduced sequence o
45	64.5	17.7	1713	21	AAB48458	Human laminin 5 po

ALIGNMENTS

RESULT 1
AAB18677
ID AAB18677 standard; peptide; 70 AA.
XX
AC AAB18677;
XX
DT 22-JAN-2001 (first entry)
XX
DE Peptide fragment from Rep protein of TGMV (amino acids 110-179).
XX
KW Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW ribosome binding region; resistance; geminivirus infection.
XX
OS Tomato golden mosaic virus.
XX
PN WO200054573-A1.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06759.
XX
PR 18-MAR-1999; 99US-0125004.
PR 09-APR-1999; 99US-0289346.
XX
PA (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI Hanley-Bowdoin L, Orozco BM, Kong L;
XX
DR WPI; 2000-618851/59.
XX
PT Transgenic plants with increased resistance to geminivirus infection
PT comprise a nucleic acid construct containing a nucleic acid sequence
PT encoding a mutant AL1 protein with a mutation in the Rb binding region
PT -

```
XX
PS   Disclosure; Page 18; 73pp; English.
XX
CC   The present sequence is derived from a geminivirus replication (Rep)
CC   protein, which is also known as AL1. AL1 binds double-stranded DNA,
CC   catalyses cleavage and ligation of single-stranded DNA, and interacts
CC   with other viral and host proteins. Mutants of the AL1 protein are used
CC   to produce transgenic plants. The mutation in AL1 is present in a
CC   ribosome binding region, and expression of mutant AL1 protein imparts
CC   increased resistance to geminivirus infection in the plant. Mutant AL1
CC   proteins are useful for producing plants having increased resistance or
CC   reduced sensitivity to a geminivirus such as tomato golden mosaic virus,
CC   tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl
CC   virus, African cassava mosaic virus, Indian cassava mosaic virus, potato
CC   yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,
CC   squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or
CC   beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             100.0%;  Score 364;  DB 21;  Length 70;
  Best Local Similarity   100.0%;  Pred. No. 3.7e-38;
  Matches   70;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT   2
AAB18687
ID   AAB18687 standard; peptide; 356 AA.
XX
AC   AAB18687;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Amino acid sequence of a geminivirus replication protein of TGMV.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 354
FT                   /note= "unspecified amino acid"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
```

```
PS   Disclosure; Page 47-48; 73pp; English.
XX
CC   The present sequence represents a geminivirus replication (Rep)
CC   protein, which is also known as AL1. AL1 binds double-stranded DNA,
CC   catalyses cleavage and ligation of single-stranded DNA, and interacts
CC   with other viral and host proteins. Mutants of the AL1 protein are used
CC   to produce transgenic plants. The mutation in AL1 is present in a
CC   ribosome binding region, and expression of mutant AL1 protein imparts
CC   increased resistance to geminivirus infection in the plant. Mutant AL1
CC   proteins are useful for producing plants having increased resistance or
CC   reduced sensitivity to a geminivirus such as tomato golden mosaic virus,
CC   tomato mottle virus, tomato yellow leaf curl virus, tomato leaf curl
CC   virus, African cassava mosaic virus, Indian cassava mosaic virus, potato
CC   yellow mosaic virus, bean golden mosaic virus, bean dwarf mosaic virus,
CC   squash leaf curl virus, Texas pepper virus, cotton leaf curl virus or
CC   beet curly top virus.
XX
SQ   Sequence   356 AA;


  Query Match             100.0%;  Score 364;  DB 21;  Length 356;
  Best Local Similarity   100.0%;  Pred. No. 3e-37;
  Matches   70;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     110 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 169

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     170 drifdktpep 179


RESULT   3
AAB18685
ID   AAB18685 standard; peptide; 70 AA.
XX
AC   AAB18685;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 10
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
```

```
PS   Claim 53; Page 46; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             97.8%;  Score 356;  DB 21;  Length 70;
  Best Local Similarity   98.6%;  Pred. No. 3.8e-37;
  Matches   69;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 tlvwgefqvagrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT   4
AAB18688
ID   AAB18688 standard; peptide; 70 AA.
XX
AC   AAB18688;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 19
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 20
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 48; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             97.0%;  Score 353;  DB 21;  Length 70;
  Best Local Similarity   97.1%;  Pred. No. 9e-37;
  Matches   68;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db       1 tlvwgefqvdgrsarggcaasndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT   5
AAB18692
ID   AAB18692 standard; peptide; 70 AA.
XX
AC   AAB18692;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 66
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 69
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
```

```
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 50; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             96.7%;  Score 352;  DB 21;  Length 70;
  Best Local Similarity   97.1%;  Pred. No. 1.2e-36;
  Matches   68;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||| || |
Db      61 drifdatpap 70


RESULT   6
AAB18684
ID   AAB18684 standard; peptide; 70 AA.
XX
AC   AAB18684;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 7
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 8
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
```

```
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             96.2%;  Score 350;  DB 21;  Length 70;
  Best Local Similarity   97.1%;  Pred. No. 2.1e-36;
  Matches   68;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 tlvwgeaavdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT   7
AAB18690
ID   AAB18690 standard; peptide; 70 AA.
XX
AC   AAB18690;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 27
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 30
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
```

```
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             96.2%;  Score 350;  DB 21;  Length 70;
  Best Local Similarity   97.1%;  Pred. No. 2.1e-36;
  Matches   68;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          ||||||||||||||||||||||||| || ||||||||||||||||||||||||||||||
Db      1 tlvwgefqvdgrsarggcqtsndaaaaalaasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   8
AAB18678
ID   AAB18678 standard; peptide; 70 AA.
XX
AC   AAB18678;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 12
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 13
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 15
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 42-43; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             95.9%;  Score 349;  DB 21;  Length 70;
  Best Local Similarity   95.7%;  Pred. No. 2.9e-36;
  Matches   67;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |||||||||| :| ||||||||||||||||||||||||||||||||||||||||||||||
Db      1 tlvwgefqvdgaaaaggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy     61 DRIFDKTPEP 70
          ||||||||||
Db     61 drifdktpep 70


RESULT   9
AAB18686
ID   AAB18686 standard; peptide; 70 AA.
XX
AC   AAB18686;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 24
FT                   /note= "wild type residue replaced with Leu"
FT   Misc-difference 25
FT                   /note= "wild type residue replaced with Leu"
FT   Misc-difference 26
FT                   /note= "wild type residue replaced with Leu"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
```

```
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 46; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             95.9%;  Score 349;  DB 21;  Length 70;
  Best Local Similarity   95.7%;  Pred. No. 2.9e-36;
  Matches   67;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndlllealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  10
AAB18689
ID   AAB18689 standard; peptide; 70 AA.
XX
AC   AAB18689;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 22
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 23
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 48-49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             95.6%;  Score 348;  DB 21;  Length 70;
  Best Local Similarity   97.1%;  Pred. No. 3.8e-36;
  Matches   68;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsaaaaaealnasskeealqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  11
AAB18680
ID   AAB18680 standard; peptide; 70 AA.
XX
AC   AAB18680;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 42
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 43
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 44
```

```
FT                  /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 43-44; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             95.1%;  Score 346;  DB 21;  Length 70;
  Best Local Similarity   95.7%;  Pred. No. 6.8e-36;
  Matches   67;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiiaaaipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  12
AAB18691
ID   AAB18691 standard; peptide; 70 AA.
XX
AC   AAB18691;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 34
FT                  /note= "wild type residue replaced with Ala"
FT   Misc-difference 35
FT                  /note= "wild type residue replaced with Ala"
FT   Misc-difference 36
FT                  /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Disclosure; Page 49; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             95.1%;  Score 346;  DB 21;  Length 70;
  Best Local Similarity   95.7%;  Pred. No. 6.8e-36;
  Matches   67;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnassaaaalqiirekipekylfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  13
AAB18681
ID   AAB18681 standard; peptide; 70 AA.
XX
AC   AAB18681;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
```

```
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 47
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 48
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 49
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 52; Page 44; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             94.2%;  Score 343;  DB 21;  Length 70;
  Best Local Similarity   95.7%;  Pred. No. 1.6e-35;
  Matches   67;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipaaalfqfhnlnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


RESULT  14
AAB18683
ID   AAB18683 standard; peptide; 70 AA.
XX
AC   AAB18683;
XX
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 59
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 61
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 62
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match             94.0%;  Score 342;  DB 21;  Length 70;
  Best Local Similarity   95.7%;  Pred. No. 2.2e-35;
  Matches   67;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfqfhnlnsal 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 aaifdktpep 70


RESULT  15
AAB18682
ID   AAB18682 standard; peptide; 70 AA.
XX
AC   AAB18682;
XX
```

```
DT   22-JAN-2001  (first entry)
XX
DE   Mutant peptide derived from amino acids 110-179 of Rep (AL1) protein.
XX
KW   Geminivirus; replication protein; Rep protein; AL1; transgenic plant;
KW   ribosome binding region; resistance; geminivirus infection.
XX
OS   Synthetic.
OS   Tomato golden mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 52
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 54
FT                   /note= "wild type residue replaced with Ala"
FT   Misc-difference 55
FT                   /note= "wild type residue replaced with Ala"
XX
PN   WO200054573-A1.
XX
PD   21-SEP-2000.
XX
PF   15-MAR-2000; 2000WO-US06759.
XX
PR   18-MAR-1999;   99US-0125004.
PR   09-APR-1999;   99US-0289346.
XX
PA   (UYNC-) UNIV NORTH CAROLINA STATE.
XX
PI   Hanley-Bowdoin L,  Orozco BM,  Kong L;
XX
DR   WPI; 2000-618851/59.
XX
PT   Transgenic plants with increased resistance to geminivirus infection
PT   comprise a nucleic acid construct containing a nucleic acid sequence
PT   encoding a mutant AL1 protein with a mutation in the Rb binding region
PT   -
XX
PS   Claim 53; Page 44-45; 73pp; English.
XX
CC   The present sequence represents a mutant peptide, derived from a
CC   geminivirus replication (Rep) protein, also known as AL1. AL1 binds
CC   double-stranded DNA, catalyses cleavage and ligation of single-stranded
CC   DNA, and interacts with other viral and host proteins. Mutants of the AL1
CC   protein are used to produce transgenic plants. The mutation in AL1 is
CC   present in a ribosome binding region, and expression of mutant AL1
CC   protein imparts increased resistance to geminivirus infection in the
CC   plant. Mutant AL1 proteins are useful for producing plants having
CC   increased resistance or reduced sensitivity to a geminivirus such as
CC   tomato golden mosaic virus, tomato mottle virus, tomato yellow leaf curl
CC   virus, tomato leaf curl virus, African cassava mosaic virus, Indian
CC   cassava mosaic virus, potato yellow mosaic virus, bean golden mosaic
CC   virus, bean dwarf mosaic virus, squash leaf curl virus, Texas pepper
CC   virus, cotton leaf curl virus or beet curly top virus.
XX
SQ   Sequence   70 AA;


  Query Match           93.4%;  Score 340;  DB 21;  Length 70;
  Best Local Similarity  95.7%;  Pred. No. 3.9e-35;
  Matches   67;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||||||||||||||||| |  |||||
Db       1 tlvwgefqvdgrsarggcqtsndaaaealnasskeealqiirekipekylfafaalnsnl 60

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db      61 drifdktpep 70


Search completed: October 10, 2002, 01:47:17
```

Job time: 526 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 10, 2002, 02:01:59 ; Search time 118.38 Seconds
(without alignments)
102.295 Million cell updates/sec

Title: US-09-289-346A-3
Perfect score: 358
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	293	81.8	351	12	Q91R10	Q91r10 tomato seve
2	291	81.3	352	12	Q9E000	Q9e000 tomato rugo
3	276	77.1	226	12	O09727	O09727 leonurus mo
4	271	75.7	314	12	Q9ELT8	Q9elt8 sweet potat
5	271	75.7	364	12	Q9QS55	Q9qs55 sweet potat
6	264	73.7	226	12	Q9WHF6	Q9whf6 tomato mild
7	263	73.5	359	12	Q91M88	Q91m88 tobacco lea
8	262	73.2	361	12	Q67574	Q67574 bean golden
9	260	72.6	149	12	P88975	P88975 macroptiliu
10	260	72.6	225	12	Q9QDB1	Q9qdb1 cowpea gold
11	260	72.6	233	12	Q9YLA4	Q9yla4 macroptiliu
12	259	72.3	185	12	Q98693	Q98693 sida golden
13	251	70.1	190	12	Q9Z089	Q9z089 tobacco lea
14	251	70.1	190	12	Q9Z084	Q9z084 tobacco lea
15	250	69.8	190	12	Q9W827	Q9w827 tobacco lea
16	250	69.8	208	12	Q9Z0C4	Q9z0c4 tobacco lea
17	250	69.8	208	12	Q9Z0C0	Q9z0c0 tobacco lea
18	250	69.8	208	12	Q9Z0B8	Q9z0b8 tobacco lea
19	249	69.6	208	12	Q9Z0A0	Q9z0a0 tobacco lea
20	248	69.3	208	12	Q9Z0C6	Q9z0c6 tobacco lea
21	247	69.0	362	12	O56816	O56816 chayote mos
22	246	68.7	208	12	Q9Z0B6	Q9z0b6 tobacco lea
23	245	68.4	203	12	Q9Z083	Q9z083 tobacco lea
24	245	68.4	363	12	O73577	O73577 cotton leaf
25	244	68.2	362	12	Q9IN48	Q9in48 okra enatio
26	243	67.9	234	12	O39180	O39180 geminivirid
27	242	67.6	231	12	Q96620	Q96620 african tom
28	242	67.6	353	12	O72692	O72692 beet curly
29	242	67.6	354	12	Q65438	Q65438 beet curly
30	242	67.6	359	12	Q9YZV4	Q9yzv4 tomato yell
31	242	67.6	359	12	Q9YZV2	Q9yzv2 tomato yell
32	242	67.6	359	12	Q9YUX7	Q9yux7 tomato yell
33	242	67.6	359	12	Q9YL27	Q9yl27 tomato yell
34	241	67.3	358	12	Q65418	Q65418 beet curly
35	241	67.3	359	12	Q88942	Q88942 tomato yell
36	240	67.0	359	12	Q91B86	Q91b86 ageratum ye
37	240	67.0	360	12	Q9DXI0	Q9dxi0 ageratum ye
38	239	66.8	307	12	Q9IET7	Q9iet7 cotton leaf
39	239	66.8	361	12	O72723	O72723 cotton leaf
40	238	66.5	190	12	Q9Z0A7	Q9z0a7 tobacco lea
41	238	66.5	231	12	Q67609	Q67609 tomato leaf
42	238	66.5	348	12	Q911W5	Q911w5 macroptiliu
43	237	66.2	190	12	Q9Z086	Q9z086 tobacco lea
44	234	65.4	363	12	O72705	O72705 cotton leaf
45	234	65.4	363	12	O72719	O72719 cotton leaf

ALIGNMENTS

RESULT 1
Q91R10
ID Q91R10 PRELIMINARY; PRT; 351 AA.
AC Q91R10;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE REP PROTEIN.
GN AC1.
OS Tomato severe rugose virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=158463;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MINAS GERAIS;
RA Rezende W.L., Goulart L.R., Parreira K.S., Figueiredo J.E.F.;
RT "The full-length DNA-A nucleotide sequence of a novel tomato-infecting
RT begomovirus, Tomato severe rugose virus, in Brazil.";
RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY029750; AAK50357.1; -.
SQ SEQUENCE 351 AA; 40122 MW; 87F937A4F873B6CF CRC64;

Query Match 81.8%; Score 293; DB 12; Length 351;
Best Local Similarity 78.6%; Pred. No. 3.5e-26;
Matches 55; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:||||||||||||:||||| || | |   ||||||||:|||:||||||||||
Db     111 TIEWGEFQIDGRSARGGCQTANDAAAEALNAPSKDVALQIIREKLPEKFLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
             |||| : |||
Db     171 DRIFARAPEP 180
```

RESULT 2
Q9E000
ID Q9E000 PRELIMINARY; PRT; 352 AA.

```
AC   Q9E000;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN.
GN   AC1.
OS   Tomato rugose mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=134599;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Fernandes J.J., Fontes E.P.B., Brommonschenkel S.H., Carvalho M.G.,
RA   Zambolim E.M., Zerbini F.M.;
RT   "Molecular Cloning and Characterization of Tomato rugose mosaic virus
RT   (TRMV), a Begomovirus Isolated from Tomatoes at Triangulo Mineiro,
RT   Minas Gerais, Brazil.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF291705; AAG15546.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   352 AA;  40012 MW;  47CD55838E24D613 CRC64;


  Query Match             81.3%;  Score 291;  DB 12;  Length 352;
  Best Local Similarity   78.6%;  Pred. No. 6.1e-26;
  Matches   55;  Conservative    6;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:||||||||||:||||| || | |   ||||||||:|||:||||||||||
Db     111 TIEWGEFQIDGRSARGGCQTANDAAAEALNAPSKDIALQIIREKMPEKFLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| : |||
Db     171 DRIFARAPEP 180


RESULT   3
O09727
ID   O09727      PRELIMINARY;      PRT;   226 AA.
AC   O09727;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   REP.
OS   Leonurus mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=58177;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LEMV- BRAZIL 1;
RA   Faria J.C., Maxwell D.P.;
RT   "Variability in geminivirus associated with Phaseolus vulgaris in
RT   Brazil.";
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U92532; AAB51157.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     226     226
SQ   SEQUENCE   226 AA;  25617 MW;  73CDB6E766083FC5 CRC64;


  Query Match             77.1%;  Score 276;  DB 12;  Length 226;
  Best Local Similarity   77.1%;  Pred. No. 2.1e-24;
  Matches   54;  Conservative    4;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||||||||||:||| || ||||| || |      |||||:||:||||||||||||||
Db     111 TVEWGEFQVDGRSSRGGQQTVNDAAAEALNAPDKRTALQIIKEKLPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180


RESULT   4
Q9ELT8
ID   Q9ELT8      PRELIMINARY;      PRT;   314 AA.
AC   Q9ELT8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION ASSOCIATION PROTEIN.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF288227; AAG01006.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   314 AA;  35153 MW;  6B6220613046943F CRC64;


  Query Match             75.7%;  Score 271;  DB 12;  Length 314;
  Best Local Similarity   79.4%;  Pred. No. 1.2e-23;
  Matches   54;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||||||||||| ||:||||| || | |  |||||||||:|||||||||| |||
Db     110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy      61 DRIFDKTP 68
           ||||   |
Db     170 DRIFSPPP 177


RESULT   5
Q9QS55
ID   Q9QS55      PRELIMINARY;      PRT;   364 AA.
AC   Q9QS55;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION INITIATION PROTEIN AC1.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A.;
RT   "Cloning of a DNA-A-like genomic component of sweet potato leaf curl
RT   virus : nucleotide sequence and phylogenetic relationships.";
RL   Online Publication.
```

```
DR   EMBL; AF104036; AAD47173.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   364 AA;  40680 MW;  5F79752431A09D6E CRC64;


  Query Match             75.7%;  Score 271;  DB 12;  Length 364;
  Best Local Similarity   79.4%;  Pred. No. 1.4e-23;
  Matches   54;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||||||||||| ||:||||| || | |  |||||||||:||||||||||| |||
Db     110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy      61 DRIFDKTP 68
           ||||   |
Db     170 DRIFSPPP 177


RESULT   6
Q9WHF6
ID   Q9WHF6      PRELIMINARY;      PRT;   226 AA.
AC   Q9WHF6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   tomato mild mottle geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX   NCBI_TaxID=92943;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HN96-H5KW;
RA   Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA   Maxwell D.P.;
RT   "Molecular characterization and DNA-based detection methods for
RT   vegetable-infecting geminiviruses in Central America.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF131071; AAD33471.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     226    226
SQ   SEQUENCE   226 AA;  25941 MW;  2EA4116712871A23 CRC64;


  Query Match             73.7%;  Score 264;  DB 12;  Length 226;
  Best Local Similarity   70.0%;  Pred. No. 5.3e-23;
  Matches   49;  Conservative   12;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| ||:||||| || |||   |::||:||:|:::|||:|||:|||
Db     111 TIEWGEFQIDGRSARGGQQTANDAAAEALNASSKEEAMRIIKEKLPKEFLFQYHNLSSNL 170

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180


RESULT   7
Q91M88
ID   Q91M88      PRELIMINARY;      PRT;   359 AA.
AC   Q91M88;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS FROM ZIMBABWE, COMPLETE GENOME.
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Paximadis M., Rey C.;
RT   "Genome organization of tobacco leaf curl virus from Zimbabwe, a
RT   monopartite begomovirus.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF350330; AAK69688.1; -.
SQ   SEQUENCE   359 AA;  40722 MW;  349CF53FE972085D CRC64;


  Query Match             73.5%;  Score 263;  DB 12;  Length 359;
  Best Local Similarity   70.0%;  Pred. No. 1.2e-22;
  Matches   49;  Conservative    8;  Mismatches   13;  Indels    0;  Gaps.   0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           || |||||||||||||||| :||| | || | || ||: |||||:|: ::||:||| :||
Db     110 TLEWGEFQVDGRSARGGCQNANDACADALNAGSAEAAMAIIREKLPKDFIFQYHNLKTNL 169

Qy      61 DRIFDKTPEP 70
           ||||   |||
Db     170 DRIFQAPPEP 179


RESULT   8
Q67574
ID   Q67574      PRELIMINARY;      PRT;   361 AA.
AC   Q67574;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE  REPLICATIVE PROTEIN.
GN   AL1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,
RA   Maxwell D.P., Russell D.R.;
RT   "Cloning of the complete DNA genomes of four bean-infecting
RT   geminiviruses and determining their infectivity by electric discharge
RT   particle acceleration.";
RL   Phytopathology 81:980-985(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,
RA   Morales F.J., Maxwell D.P.;
RT   "Differentiation of bean-infecting geminiviruses by nucleic acid
RT   hybridization probes and aspects of bean golden mosaic in Brazil.";
RL   Plant Dis. 75:336-342(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RT   "Genetic diversity in geminiviruses causing bean golden mosaic
RT   disease: The nucleotide sequence of the infectious cloned DNA
RT   components of a Brazilian isolate of bean golden mosaic geminivirus.";
RL   Submitted (MAY-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; M88686; AAA46312.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
SQ   SEQUENCE   361 AA;  41041 MW;  0094C7ACAF06B788 CRC64;


  Query Match             73.2%;  Score 262;  DB 12;  Length 361;
  Best Local Similarity   73.1%;  Pred. No. 1.6e-22;
  Matches   49;  Conservative    9;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLDRI 63
           || |||||||||||| ||:|||||: || |||   |:|||:|||:|||:|||:|||:||||||
```

```
Db    113 WGHFQVDGRSARGGQQTANDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy     64 FDKTPEP 70
          | | |:|
Db    173 FTKAPDP 179


RESULT   9
P88975
ID   P88975      PRELIMINARY;      PRT;   149 AA.
AC   P88975;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   AC1.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICAN;
RA   Roye M.E.;
RL   Thesis (1996), Biochemistry, University of the West Indies, Jamaica.
DR   EMBL; U75278; AAB36919.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     149    149
SQ   SEQUENCE   149 AA;  16785 MW;  E4CF5EED4C9CD508 CRC64;


  Query Match             72.6%;  Score 260;  DB 12;  Length 149;
  Best Local Similarity   68.6%;  Pred. No. 9.7e-23;
  Matches   48;  Conservative   12;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          |: || ||:|||||||| |||||||| || : :  ||::|:::||:|||:|||:|||:|||
Db     52 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 111

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    112 DRIFMKDPEP 121


RESULT  10
Q9QDB1
ID   Q9QDB1      PRELIMINARY;      PRT;   225 AA.
AC   Q9QDB1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   cowpea golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=69263;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CGMV-BR;
RA   Faria J.C.;
RT   "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT   Brazil.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF188708; AAF06318.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     225    225
SQ   SEQUENCE   225 AA;  25766 MW;  1089CB6BD8D15B5D CRC64;


  Query Match             72.6%;  Score 260;  DB 12;  Length 225;
  Best Local Similarity   73.1%;  Pred. No. 1.6e-22;
  Matches   49;  Conservative    8;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLDRI 63
          || ||:|||||||| || |||| : || |||   |:|||:||:|||:|||:|||:||||||
Db    113 WGHFQIDGRSARGGQQTINDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy     64 FDKTPEP 70
          | | |||
Db    173 FKKPPEP 179


RESULT  11
Q9YLA4
ID   Q9YLA4      PRELIMINARY;      PRT;   233 AA.
AC   Q9YLA4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E.;
RT   "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
RT   from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E., McLaughlin W.A., Maxwell D.P.;
RT   "Molecular characterization of two distinct geminiviruses infecting M.
RT   lathyroides from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF098940; AAD17850.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER     233    233
SQ   SEQUENCE   233 AA;  26356 MW;  AA490AF4D2166A02 CRC64;


  Query Match             72.6%;  Score 260;  DB 12;  Length 233;
  Best Local Similarity   68.6%;  Pred. No. 1.6e-22;
  Matches   48;  Conservative   12;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          |: || ||:|||||||| |||||||| || : :  ||::|:::||:|||:|||:|||:|||
Db    110 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 169

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    170 DRIFMKDPEP 179


RESULT  12
Q98693
ID   Q98693      PRELIMINARY;      PRT;   185 AA.
AC   Q98693;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   AC1.
OS   sida golden mosaic virus.
```

```
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51034;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA;
RA   Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RT   "Genetic Diversity among geminiviruses associated with the weed
RT   species Sida spp, Macroptilium lathyroides, and Wissadula amplissima
RT   from Jamaica.";
RL   Plant Dis. 81:1251-1258(1997).
DR   EMBL; U67926; AAB97865.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     185    185
SQ   SEQUENCE   185 AA;  20975 MW;  3913850A025A5EE1 CRC64;


  Query Match             72.3%;  Score 259;  DB 12;  Length 185;
  Best Local Similarity   68.6%;  Pred. No. 1.6e-22;
  Matches   48;  Conservative   11;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:|||||||| ||:||||| || : :   ||:|||||:||:||||:|||:||:
Db      89 TIEWGVFQIDGRSARGGQQTANDAAAEALNSGTKEDALKIIREKLPERYLFQYHNLSSNI 148

Qy      61 DRIFDKTPEP 70
             |||| | |||
Db     149 DRIFSKPPEP 158


RESULT  13
Q9Z089
ID   Q9Z089      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z089;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA3-1, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001315; BAA34033.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21432 MW;  AAC093D1D1610FAD CRC64;


  Query Match             70.1%;  Score 251;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 1.4e-21;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           || || ||:|||||||||| :||| | || ||| | || |||||:|: ::||:|||||||
Db      85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  14
Q9Z084
ID   Q9Z084      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z084;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE YOKOHAMA5-2, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA5;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001318; BAA34039.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  AAC1C2943E3F01AD CRC64;


  Query Match             70.1%;  Score 251;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 1.4e-21;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           || || ||:|||||||||| :||| | || ||| | || |||||:|: ::||:|||||||
Db      85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  15
Q9W827
ID   Q9W827      PRELIMINARY;      PRT;   190 AA.
AC   Q9W827;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   TOBACCO LEAF CURL VIRUS C1 AND C4 GENES, CLONE GORA1-1, PARTIAL AND
DE   COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GORA1;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001303; BAA34010.1; -.
DR   InterPro; IPR001191; Gemini_AL1.
DR   Pfam; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   ProDom; PD000736; Gemini_AL1; 1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  93C3742A8EBDB7EB CRC64;
```

Query Match 69.8%; Score 250; DB 12; Length 190;
Best Local Similarity 58.8%; Pred. No. 1.9e-21;
Matches 50; Conservative 8; Mismatches 11; Indels 16; Gaps 1;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           || || ||||||||||||| :||| | || |:| | || |||||:|: ::||:|||||||
Db      85 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFSNSSFDQVPE 169
```

Search completed: October 10, 2002, 02:01:59
Job time: 1092 sec